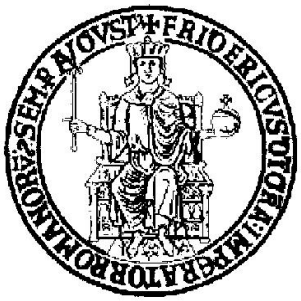
UNIVERSITÀ DEGLI STUDI DI NAPOLI
"FEDERICO II"

Degree of Doctor of Philosophy



PhD Thesis

Mathematical Modeling and Non-linear Analysis of Synthetic and Endogeneous Gene Regulatory Networks

Guide:

Ing. Diego di Bernardo

Prof. Mario di Bernardo

Candidate:

Immacolata Garzilli

Years 2010/2013

Imagination is more important than knowledge.
For knowledge is limited, whereas imagination embraces the entire world,
stimulating progress, giving birth to evolution.
A. Einstein

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Complex systems are open self-organizing structures able to modify their behaviour and properties due to the interaction with the external environment. These modifications depend on the structure and conformation of the systems themselves, but also on the available resources or on presence of other interacting complex systems. Interactions with environment and other systems are mediated by feedbacks. Examples of complex systems are living organisms with their features depending on their own genetic code. Their function and development depend on the molecular species, such as coding and non-coding genes encoded in their inherited genome, but also on the interactions between them: systems of interacting genes are called *gene regulatory networks*.

Gene regulatory networks are characterised by positive or negative regulatory interactions between transcription factors and their target genes:

their composition allows the network to perform different functions in different contexts, thus conferring specific dynamical properties to the expressed genes. Considering the single components (transcription factors and regulated genes) it is possible to use mathematical models to evaluate the behaviour of the whole system, thus revealing its function and predicting its behaviour as a result of endogeneous and exogeneous stimuli. One way to characterize the single components in order to predict the function of the complete system is to describe the single parts using differential equations and looking for a solution able to reproduce the measured response of the system.

This kind of analysis is typical of the Dynamical Systems Theory and it has been widely applied to physical, engineered, economical, social and artificial systems and recently to biological ones. What it is known from Dynamical Systems Theory is that when the single components act following nonlinear laws, the behaviour of the complex system cannot be deduced as the simple sum of its parts, but unexpected properties can arise, thus leading to what are called *emergent properties*: they are known to appear when a number of simple entities (agents) operates in an environment, giving rise to more complex behaviors, as a collective, in an autonomous way. For example, synchronization of coupled autonomous can be seen as an emergent property of the self-organization between single agents.

In this Thesis, I propose to study a subset of gene regulatory network

topologies, thus explaining the function of common regulatory motifs observed at molecular level, highligthing the appeareance of unexpected behaviours due to the intrinsic nonlinearities and also revealing the versatility of some of these networks in performing different actions, simply varying the behaviour of a single component. First, I will focus on some synthetic gene networks constructed *ad-hoc* so that the individual biological parts are well-known and thus they can be easily characterized: this is a classical approach used in the field of Synthetic Biology. Then, I will theoretically describe an endogeneous gene network: in this case, the complex system under investigation is first, modelled in a single cell and then as a "network of networks" across multiple cells. The investigation of endogeneous gene networks is central in the field of Systems Biology.

1.1 Systems and Synthetic Biology

In this work, approaches commonly used in the fields of Synthetic and Systems Biology will be applied. These disciplines are related to the description of biological systems in order to uncover the basic properties using tools from Mathematics, Engineering and Computer Science. In particular, while Systems Biology aims at elucidating endogeneous systems starting from esperimental evidence, Synthetic Biology describes the molecular mechanisms of gene regulatory network through the rational construction of simplified regulatory motifs [33].

Systems Biology is the study of biological systems using quantitative measurements of the behavior of groups of interacting components: systematic measurement technologies such as genomics, bioinformatics and proteomics, and mathematical and computational models to describe and predict dynamical behavior. Data coming from experiments or simple observations can be analysed thus inferring the network topology [74].

So, while Systems Biology tries to deduce the components of the biological networks of interest validating hyphoteses based on the observable behaviours, Synthetic Biology is based on the idea that reproducing synthetically a network can lead to the comprehension of basic properties at molecular level. Thus, biological networks are constructed *de novo* with known components and with known and predefined interactions, thus making possible the analysis of a complete circuit starting from its single parts. Making Biology "synthetic" is a way to test what kind of interactions (for example in terms of quality, acting as inhibitors or activators or promoters of degradation, strength, timescales) are needed to lead the cell to develop particular functions and what of these properties can be modified to eventually improve these features. Understanding the relationship between topology and dynamics of transcriptional regulatory networks in mammalian cells is essential to elucidate the biology of complex regulatory and signaling pathways. Different circuits acting different functions (monostable, bistable, oscillating systems and inducible ones are some examples) have been designed and implemented

from bacteria to yeast and mammalian cells [52], [139], [27]. If on one hand, synthetic circuits suggest what are the molecular mechanisms needed to operate in a certain manner, they can also be used to perform specific functions: this is the case of the many applications in which synthetic networks have been built and integrated in cells in order to be used for treatment of diseases or to function as biosensors, to produce energy or transform compounds ([79], [72]). Another effort taking place in Synthetic Biology is the building of a minimal cell able to reproduce specific cell functions [132].

1.2 Motivation and Thesis outline

The main aim of this thesis is to provide a mathematical characterization of some very common motifs in gene regulation, analyzing their dynamical properties starting from synthetic networks and then moving to endogeneous ones. The networks here considered can exhibit a variety of dynamical behaviours, such as monostable, bistable or oscillatory dynamics. Moreover, biological "parts" can perform different functions depending on the context, as in the case of microRNAs. Moreover, besides the description of the complexity of regulation inside a single cell, feedback regulation happens also among cells, thus linking the networks into a more complex multicellular dynamical system.

In Chapter 2, some preliminary concepts on the biology underlying the systems under investigation are given as well as some mathematical tools for building dynamical models of them.

In Chapter 3, the first synthetic network of interest is considered: an inducible Positive Feedback Loop (PFL) integrated in a mammalian cell line is described and modeled, thus comparing its dynamical features to the properties of an *ad hoc* designed circuit in which no feedback is considered (NOPFL).

In Chapter 4, the characterization of a more complex synthetic network is provided: a negative feedback loop mediated by a microRNA is added to the previous PFL, thus generating a Positive Negative Feedback Loop (PNFL). The dynamical properties of this new circuit are studied and compared to the PFL network, thus revealing the role of the microRNA in speeding up the network dynamics. An experimental validation is also provided, thanks to the use of a microfluidic device designed for handling mammalian cells (courtesy of Prof Jeff Hasty at the University of California, San Diego).

In Chapter 5, the role of the microRNA in suppressing biological noise is investigated. By comparing PFL to PNFL it is possible to study the variation in gene espression across genetically identical clonal cells. We observed

that clones expressing the PNFL have a reduced variation in gene expression when compared to PFL cells. The experimental observation is confirmed and explained by dynamical and stochastic models of both the PFL and the PNFL networks.

In Chapter 6, a literature review on the Notch-Delta Pathway is provided. The Notch and Delta pathway is a prototype signalling pathway enabling "communication" among different cells.

In Chapter 7, a mathematical description of Notch-Delta Pathway is presented. In particular, when cells are communicating through Notch and Delta, "patterns" of gene expression can emerge in the tissue. The ability of this signaling pathways to give rise to these inhomogeneous configurations of gene expression is mathematically investigated.

In Chapter 8, I analyse a more peculiar interaction between Notch-Delta and one of its effector genes: the Hes1 transcription factor, which exhibits autonomous oscillations in expression in specific cell types and tissues. In this case the behaviour of a tissue of autonomously oscillating cells is considered and studied depending on the properties of the single components, such as oscillator's features or communication channel characteristics.

The same kind of system is considered in Chapter 9, in which a study on the synchroneous behaviour of the different oscillators is perfomed, thus also contributing to a new hypothesis on the vertebrate segmentation process, usually explained through the use of the classical Clock and Wavefront model.

To conclude, Chapter 10 provides a summary of the main results.

Furthermore, in Appendix A, more details on the experimental data and material and methods are provided, together with the main algorithms used for the numerical investigation of the models and for processing experimental data.

All the synthetic circuits considered in this Thesis have been implemented in vivo by Dr. Velia Siciliano at TIGEM (Dr. Diego di Bernardo - Synthetic and Systems Biology Lab).

Chapter 2

Preliminary Concepts

In this chapter some general concepts are provided on the meaning of biological terms and the physical mechanisms involved in gene regulatory networks and in network motifs. Moreover, I present an overview of the different methods to mathematically model and analyse gene regulatory networks.

2.1 Gene regulation and network motifs

Gene expression is consequence of the central dogma of Molecular Biology: genes are perpetuated as sequences of nucleic acid, but function by being expressed in the form of non-coding RNAs and proteins or for protein coding genes. *Transcription* and *translation* are the two steps responsible for their conversion from one form to the other. Starting from the part of the DNA

double helix encoding the gene of interest, transcription generates a messenger RNA (mRNA), a single-stranded RNA identical in sequence with one of the strands of the duplex DNA. In protein-coding genes, translation will convert the nucleotide sequence of mRNA into the sequence of amino acids comprising a protein. This two-stage process is called *gene expression*: it is a complex process regulated at several stages in the synthesis of proteins [18]. Some proteins are structural and will accumulate at the cell-wall or within the cell to give it particular properties. Other proteins can be enzymes that catalyse certain reactions. A large group of proteins have an important role in the regulation of the genes, known as *transcription factors* (TF): these proteins are able to bind to specific DNA sequences, thereby promoting or blocking (eventually with other proteins in a complex) the recruitment of RNA polymerase (the enzyme that performs the transcription of genetic information from DNA to RNA) to specific genes. So, genes are able to regulate each other, thus modifying their expression and giving rise to large regulatory networks.

There is a great variety of regulatory systems and organisms in which feedback loops have been identified. The transcriptional interactions between TF and target genes can be positive or negative, in the sense that a TF can either activate or inhibit the transcription of its target genes, or even its own transcription thus giving rise to positive or negative feedback loops.

In negative regulation, a TF binds to the promoter of a target gene (a

region of DNA that facilitates the transcription of a particular gene), and decreases the mRNA transcription, Figure 2.1A. In positive regulation, a transcription factor is required to bind the promoter in order to increase the mRNA transcription rate [6] (Figure 2.1 B). Several other steps in the gene expression process may be modulated.

When a TF binds its own promoter, it can give rise to a Negative Feedback Loop (NFL) if it behaves as an inhibitor of transcription, or to a Positive Feedback Loop (PFL) if it behaves as an activator of the transcription.

The duality between positive and negative feedbacks has been predicted in a biological setting [121], and it is a well established concept in "control engineering", a branch of engineering which deals with the design of automated mechanisms to control a variable of interest (the altitude of an airplane, or more simply, the temperature of a room via thermostat) [5]. Specifically, the negative feedback loop is a classic control engineering approach useful to regulate the variable of interest and at the same time to speed up the response times of the system, thus quickly achieving a desired value. Positive feedback loops, on the contrary, can slow down the response of a dynamical system to an external input. PFLs are used by control engineers to build "memory" elements, also known as switches, which are able to be in one of the two stable steady-states (ON or OFF), and which are robust against unwanted transient perturbations that may inadvertently switch off (or on) the system. Regulations between genes can act at different levels: indeed negative feed-

backs can act either at the level of gene transcription or by promoting degradation of the mRNAs. microRNAs (miRNAs), for example, are small non-coding RNAs able to downregulate the expression of target genes by both promoting their degradation and blocking protein translation.

In addition to feedback loops, other regulatory "motifs" are also possible:

The *Feed-Forward Loop* (FFL) is defined by a transcription factor (X) that regulates a second transcription factor (Y), such that both X and Y jointly bind a common target (Z). Since each of the regulatory interactions may either be positive or negative there can be eight different types of FFL motifs. Two of the most frequently found in gene regulatory networks are the coherent type 1 showed in Figure 2.1 (C)(where all interactions are positive) and the incoherent type (X activates Z and also activates Y which represses Z). In most of the cases the FFL can implement either an AND gate (X and Y are required for Z activation) or OR gate (either X or Y are sufficient for Z activation) but other input functions are also possible [6].

The *regulatory chain* consists of chains of three or more transcription factors in which one regulator binds the promoter of a second regulator, and in turn the second binds the promoter of a third regulator and so forth.

An *auto-regulation* or *feedback motif* consists of a transcription factor that binds its own promoter. The regulation can be positive or negative depending on the properties desired, as discussed in details above.

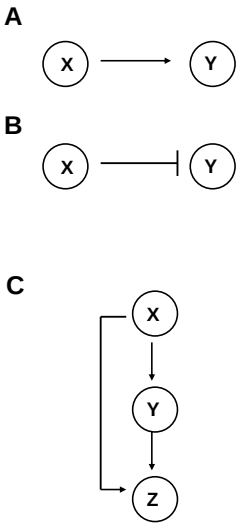


Figure 2.1: **Gene interactions**. (A) Directed edges with an arrow end represent activation. (B) Directed edges with a dash end represents inhibition. (C) Schematic representation of a coherent feed-forward Loop.

A *multi-component loop* motif consists of a regulatory circuit involving two or more factors. The closed loop structure provides the capacity for the feedback control and offers the potential to produce bistable systems, which switch between two alternative states or oscillatory dynamics.

The *single input* motif contains a single regulator that binds a set of target genes. In this way the expression of the target genes is coordinated under a specific condition. In the *multiple input* motif, there is the presence of multiple regulators acting on the metabolites of the system.

By combining together these network motifs, more complex systems can be generated, like genetic switches ([52], [77], [39]) or oscillators ([99], [59], [114]).

2.2 Mathematical modeling of gene regulatory networks

Synthetic and Systems Biology are based on the interactions between Biology and Engineering, Physics and Applied Mathematics. Quantitative models are useful instruments to describe, in terms of single components a synthetic (or endogeneous) network under investigation, thus providing both a believable understanding of the network and enabling quantitative predictions of its

behaviour. A dynamical model is a formalization of the biological knowledge about a certain system, where each component of the system is described by an equation, which represents its behaviour as a function of its regulators. *In silico* predictions of the behaviour of a biological system can be used to complement in vivo experimental observations and accelerate the hypothesis generation-validation cycle of research.

In this thesis, deterministic and stochastic dynamical models are used to investigate the properties of synthetic and endogeneous systems. From a modeling perspective, biological circuits can be analysed by looking at the features of their single parts and products: promoter leakiness and strength, transcription factors activity, mRNA and protein translation rates and decays, the action of microRNAs, and more. Different models can be used to investigate different properties of the same circuit: the desired prediction of a system of interest can be in different contexts. For example, in the case of a Positive Negative Feedback Loop as described in Chapters 4 and 5, we are interested both in the analysis of its dynamical properties and on its robustness to external perturbations: in the former case a deterministic model has to be used, in the latter a stochastic one is needed.

Ideally, when modelling a biological process, the previous knowledge needs to be formalized in the chosen framework, and all the relavant information (not only concentrations and rates of events, but also spatial distribution, diffusion parameters, and so on) has to be known in order to make a maximally

accurate *in silico* replica of the system. Obviously, even for the best-studied systems, the mass of accumulated data still falls short of describing, even qualitatively, the variety of elementary processes that each molecular species engages in: consequently, assumptions are necessary.

A huge variety of mathematical formalisms has been proposed in the literature, such as directed graphs, Bayesian networks, Boolean networks and their generalizations, ordinary and partial differential equations, qualitative differential equations, stochastic equations, and rule-based formalisms (see, for example, [38], [42], [136] and references therein).

Quantitative models offer great detail in mimicking reality: moreover, rich qualitative insights on the system are possible using theoretical tools such as bifurcation and stability analysis, which, for example, indicate the precise boundaries of parameter ranges to which steady states or sustained oscillations correspond, or reveal the stability of the solutions before actually solving the dynamical equations representing the system. Quantitative models can be either deterministic or stochastic.

Deterministic models of Differential Equations (DEs) are commonly used to describe the average behaviour of a population of cells [38]. They have been shown to be viable for the analysis of synthetic networks in a great number of works (e. g. [45], [52], [77], [140], [133]). The reaction mechanism

is described by applying the law of mass action: the rate of any given elementary reaction is proportional to the product of the concentrations of the species reacting in the elementary process (reactants) [6].

When DEs are used, the cellular concentration of proteins, mRNAs and other molecules are represented by continuous time variables with the constraint that a concentration can not be negative. Differential equations can be Ordinary (ODEs) or Partial (PDEs) if also features like space, diffusion or deformation have to be taken into account [97].

The ODEs modelling approach is based on the following biological assumptions: the quantified concentrations do not vary with respect to space and they are continuous functions of time. These assumptions hold for processes evolving on long time scales in which the number of molecules of the species in the reaction volume is sufficiently large.

Usually, the functions describing transcriptional interactions are non-linear Hill functions or Michealis-Menten, the two differing for considering or not, respectively, the cooperativity of TF regulation on the gene of interest [6].

The Hill function models transcriptional interaction in the following way:

- activation: $H^+(y;k,h) = \frac{y^h}{y^h+k^h}$;
- repression: $H^-(y;k,h) = \frac{k^h}{y^h+k^h}$;
- combination of activation or repression

$$H^{+-} = H^+(y;k,h)(\cdot,+)H^-(z;k_1,h_1),$$

where $(\cdot, +)$ indicates that we can either sum or multiply the Hill functions in the case of multiple regulation, depending on the AND or OR kind of interaction [6]; y and z represent transcription factor levels, h are the Hill coefficients (pure numbers that refer to the cooperativity of the activation binding reaction) and k are the Michaelis-Menten constants, equal to the amount of transcription factor needed to reach half maximal activation (or repression). In the case of Michalis-Menten term, the formalism is identical, but the Hill coefficient h is fixed equal to 1. For a complete derivation of Hill and Michealis-Menten formalism please refer to [6].

After the mathematical "structure" of the model has been defined, the parameters describing the physical properties of the system have to be identified in order to obtain a full description of the interactions under investigation: the parameter estimation problem can be formulated from the mathematical viewpoint as a constrained optimization problem, where the goal is to minimize the objective function, defined as the error between model predictions and real data. In biological applications, the objective function usually displays a large number of local optima as measurements are strongly affected by noise. For this kind of problems, classical optimization methods, based on gradient descent from an arbitrary initial guess of the solution, can be unfeasible and show slow or absence of convergence. Hence, stochastic optimization algorithms could be more convenient in order to explore parameters' space in

a way more insensitive to initial conditions and avoid being trapped in local minima. In this thesis, algorithms based on "simulated annealing" tecnique has been used to estimate the numerical values of the parameters characterizing the models: specifically, the PottersWheel toolbox [84] implemented in MATLAB was used (details on functioning in Appendix A).

Alternatives to simulated annealing are possible, such as Genetic Algorithms (GA) [93] which provide a very flexible approach to non-linear optimization in synthetic networks [145], [140].

Finally, the validity and usefulness of the model have to be checked, that is its ability in predicting the behaviour of the biological process under investigation. To this end, ad-hoc experiments can be used both for parameter identification, or to test the model predictions [129]. If the predictive performances of the model are not satisfactory, it is necessary to refine the model (for example, by increasing the level of details or modifying its proper structure) and/or to perform new experiments.

Stochastic models can be used when the number of molecular species and consequently of reaction events decrease, so that the probabilistic nature of biological events becomes more evident, or when the most interesting features are related not to system dynamics level, but to its variability. Indeed, also in a monoclonal population very different levels of fluorescence of a reporter

protein can be observed. Moreover, the response of individual cells may be significantly different from the average population response. Population heterogeneity arises from stochasticity in molecular events. Several methods are used to study biological variability: the most common ones are those in which Master Equations are derived from a deterministic model to yield a stochiometric model.

Besides the mathematical models, analytical tools and numerical simulations are useful instruments to investigate the properties and the structure of the system: bifurcation analysis, qualitative studies of the solutions, linear stability analysis of the Jacobian matrix and analytical tecniques are essential to understand the main features of the system under investigation without solving it directly, but focusing on the elements necessary to obtain a determined behaviour.

Biological switches

As previously discussed, the interaction between positive and negative feedbacks yields to different and more complex gene regulatory network. A biological switch is a system that can present two (or more) different equilibria. Several genetic switches have been constructed in cells ([52], [77], [112]) using a combination of two or more negative feedbacks activated or deactivated by one or more external inducer molecules, as shown in Figure 2.2. Indeed,

a single autoregulatory positive feedback loop is enough to give rise to a bistable system and a transient stimulus is necessary to switch the system from one state to another.

From a mathematical point of view, a toggle-switch is a bistable system in which at least three equilibria coexists, two of them being stable and one unstable, as shown in Figure 2.3. The system is able to approach one of the two different steady states depending on the initial condition: the unstable equilibrium defines the boundaries of the bacins of attractions for the two equilibria. In order to experimentally switch on or off the system, it is possible to use an external inducer molecule, which acts as an external perturbation, by transiently modifying the dynamical structure of the system and forcing it into the desired bacin of attraction: when the solution, eventually crosses the unstable manifold, and it is in the right bacin of attraction, then the external stimulus can be removed, and the system will approach the new desired equilibrium. In the case of a single autoregulatory positive feedback loop, varying the strenght of the TF protein in promoting its own transcription makes the system behaviour either as a monostable dynamical system or a bistable system (switch): indeed, what defines the ability of a system of being a monostable or bistable system are the instrinsic features of the components besides its topological structure.

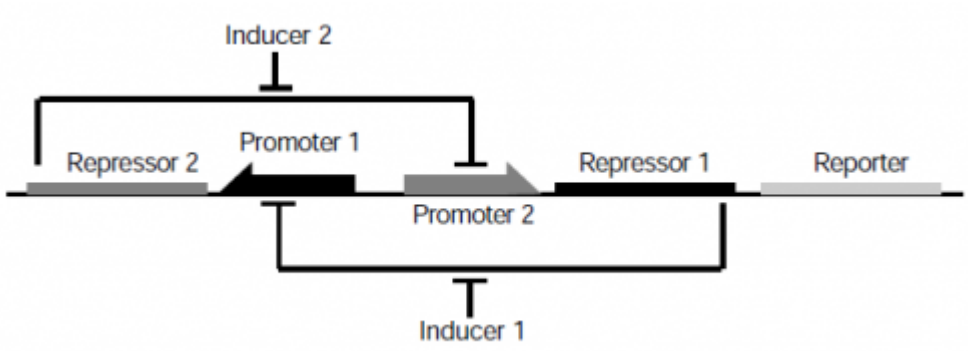


Figure 2.2: **Biological switch topology from [52].** The toggle switch is constructed from any two repressible promoters arranged in a mutually inhibitory network. It is flipped between stable states using transient chemical or thermal induction and exhibits a nearly ideal switching threshold.

Biological oscillators

A biological oscillator is a system that generates a periodic variation of a molecular species: a type of biological oscillator is the transcriptional oscillator in which the oscillation in the expression of a gene induces cyclic production of its protein product [49], [120]. In the field of Synthetic Biology, trascriptional oscillators have been constructed in systems from bacteria to mammalian cells ([45], [133], [140]). The simplest oscillator is the Goodwin oscillator [57], consisting of a single negative "delayed" feedback loop: as in the case of the bistability for a positive feedback loop, in this case the delay of the repression of the feedback on its own promoter is sufficient to sustain an oscillatory behaviour. The first synthetic oscillator was constructed in *Es-*

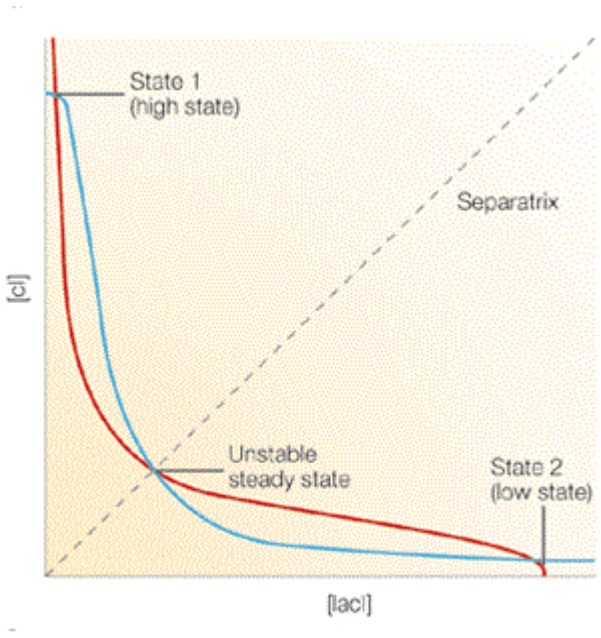


Figure 2.3: **Phase space of the bistable system described in [61].** The axes show the concentration of the two species involved in the toggle-switch. The nullclines are drawn and the intersections represent the different equilibria: the central one is the unstable and its unstable manifold is the separatrix able to distinguish between the two bacin of attraction of the stable equilibria.

cherichia Coli by [45] with the topology of a "repressilator", i.e. a chain of inhibition of three TFs as shown in Figure 2.4: in this case the oscillation is due to the odd number of negative interactions giving rise to a delayed negative feedback. From a Dynamical Systems Theory point of view, one of the tecnique to obtain an oscillatory behaviour in a dynamical system is to find parameters' values for which a Hopf bifurcation occurs. This occurs when the eigenvalues of the system dynamical matrix are complex conjugate with real part moving from negative to positive values (i.e. from a stable to an unstable equilibrium point coexisting with a stable limit cycle - from a stable to an unstable focus), as shown in Figure 2.5.

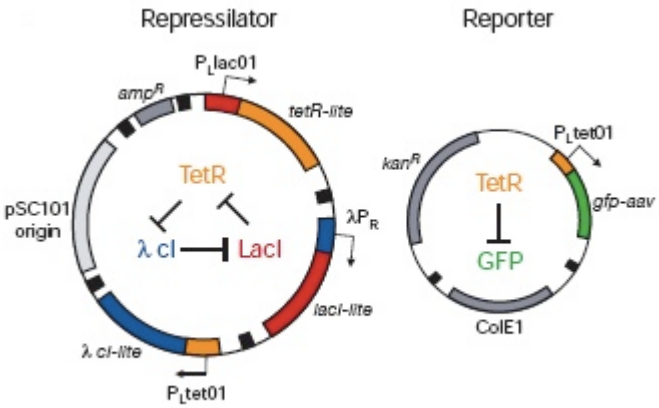


Figure 2.4: **Topology of a biological oscillator from [45].** The system is cmposed of three genes connected in a closed chain of inhibition. The circuit built in *Escherichia Coli* is able to present an oscillatory behaviour.

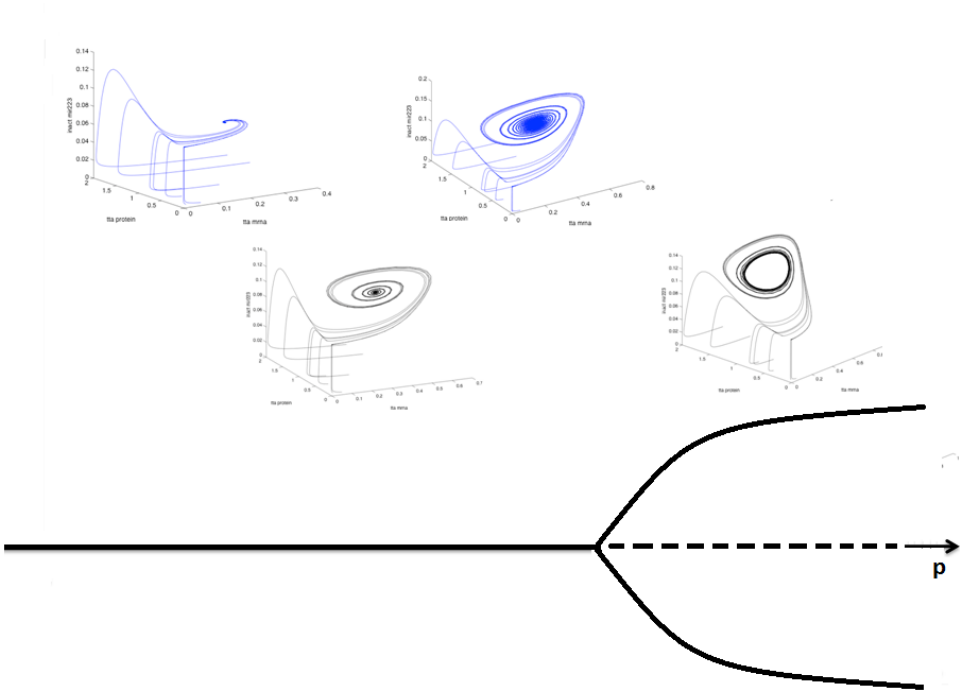


Figure 2.5: **Hopf Bifurcation.** An illustrative scheme for Hopf bifurcation is reported, in which the parameter p is considered as the varying constant. The phase space are reported for a model of three species moving p: solutions converge to a stable steady state for low values of p, while starting to converge at the limit cycle when the equilibrium is becoming unstable.

System of interacting cells

Besides the study of the dynamical properties of a molecular system in a single cell, mathematical models can be useful tools also to study the behaviour of a network of cells. Cells can be modeled as autonomous systems coupled to each other. Individual cells can communicate in different ways: electrical or chemical signals, mechanical interactions and more. Moreover, cells can communicate with each other via direct contacts (juxtacrine signaling), over short distances (paracrine signaling), or over large distances and/or scales (endocrine signaling) [48]. Some cell - cell communication mechanism requires direct cell - cell contact. Some cells can form gap junctions that connect their cytoplasm to the cytoplasm of adjacent cells. In cardiac muscle, gap junctions between adjacent cells allows for electric action potential propagation from the cardiac pacemaker region of the heart to spread and coordinately cause contraction of the heart. In the case of juxtacrine signaling, cells can communicate thanks to some signalling molecules localised on their membranes: for example, a "receptor" molecule can bind to its "ligand" molecule on the surface of the other cells.

From a mathematical and numerical point of view, it is possible to consider cells like nodes of a network and connection between them as the edges. There are different works focused on the dynamical properties of cells populations [134], [105], [61], [91]. In particular, of great interest are the theory

of biological pattern [53] or the synchronization of interacting oscillatory cells [36]. In [36] the synchronization of a population of synthetic oscillators inserted in *Escherichia Coli* is studied for the first time, describing the single clock by a DDE mathematical model and then taking into account the population behaviour by including a term for the diffusion of a coupling molecule into the medium.

The theory of synchronization and control of complex networks has been widely applied ([40], [110], [9]): a new goal of the Synthetic Biology is to use it to discover and manipulate the emergence properties of tissues.

Chapter 3

Modeling of an inducible Positive Feedback Loop stably integrated in a mammalian cell line

In this chapter, the mathematical modeling of an inducible synthetic transcriptional positive feedback loop (PFL) constructed in mammalian cells is described: the PFL is based on a *CMV-TET* promoter, responsive to the Tetracycline-controlled transactivator tTA and driving expression of the tTA protein itself and of a fluorescent reporter protein (Figure 3.1A). From an experimental point of view, a clonal population of Chinese Hamster Ovary

cells (CHO) carrying a stably integrated version of the PFL was generated in order to obtain precise quantitative measurements in a genetically homogeneous cell population. PFL clones were compared to a control clonal population of cells lacking the positive feedback loop (NOPFL), as shown in Figure 3.1B.

Two mathematical deterministic models of Ordinary Differential Equations (ODEs) were derived to predict the behaviour of the two systems (PFL and NOPFL): the model structure follows the usual "rules" as described in Chapter 2 for the modeling of transcriptional networks; the parameters, instead, were fitted using in-vivo quantitative measurements of fluorescence intensity in time, following addition of the inducer molecule (Doxycyline) able to"switch off" the PFL and NOPFL.

The two models were able to reproduce the experimental data and, interestingly, highlighted differences in the dynamic properties of the PFL versus the NOPFL networks, which are due to the intrinsic differences in the two network topologies.

My results confirmed the slowing of the dynamics of the PFL compared to the NOPFL network also in a mammalian cell system. This observation can be instrumental in better understanding the properties of natural occurring transcriptional and signaling networks.

All the results and the experimental methods were published in [126],

where also authors' contributions are declared.

3.1 Topology of the inducible positive feedback loop (PFL) and of the corresponding control network (NOPFL)

The PFL is shown in Figure 3.1A. It is based on the Tetracycline-controlled transactivator tTA, self-controlled by a *CMV-TET* promoter, responsive to the tTA itself unless the Tetracycline, or its analogous Doxycycline, is added to the medium in which cells are grown [58]. To follow the dynamics of the PFL, a destabilised yellow-green variant of the enhanced green fluorescent protein (d2EYFP), with a reported half-life of approximately two hours, is placed under the control of the same promoter. To this end, a unique cassette with an Intra Ribosomal Entry Sequence (IRES) in between of the transactivator tTA and the d2EYFP was constructed, which enables a single mRNA to encode for two different proteins (Figure 3.1A).

Clones were generated in CHO cells (details in the Appendix A or [126]). In particular, it is important to observe that more than one monoclonal population has been generated: in fact, the ability of lentiviruses to randomly integrate in the genome of the host cells, with different expression levels of the transgene according to the site of integration allowed to generate 9 PFL monoclonal populations with different level of expression of the reporter

protein (PFL numbered from 2 to 10 in Figure 5.5).

In order to study the dynamical properties of the PFL, I report the results only for one monoclonal population (PFL 7): the other clones will be useful for a more complex analysis as developed in Chapter 5.

A control network lacking the positive feedback loop (NOPFL) was generated, using the same biological "parts" as in the PFL network. As shown in Figure 3.1B, the same *CMV-TET* promoter is upstream of the d2EYFP. The tTA protein is placed under the control of a constitutive promoter, thus breaking the feedback loop. For this network a single clonal population in CHO cells was generated (NOPFL cells).

3.2 Derivation of the mathematical models for the PFL and NOPFL networks

The mathematical models for the PFL and NOPFL networks are based on sets of Ordinary Differential Equations (ODEs): indeed, measures are based on the average behaviour of a clonal population of cells. So, for each species, i.e. each mRNA and correspondent protein concentration, an equation expresses the change in concentration of the species in a given time interval, as the result of a production term and a degradation term. We assumed:

- Hill functions to model the rate of gene transcription, including basal activity to describe the leakiness of the *CMV-TET* promoter;

- linear degradation for all genes and proteins;
- linear dynamics for the translation;
- Hill functions to model the effect of the inducer (Doxycycline);
- distinct dynamics for the unfolded (inactive) and folded (active) forms of the reporter protein (d2EYFP).

The last assumption was introduced in order to take into account d2EYFP maturation time needed for correct protein folding [140]. Thus, two differential equations are considered for protein maturation as in [140]: one for the translation of mRNA to the unfolded d2EYFP protein, and one for the folded protein d2EYFP.

Letting x_1 be the *tTA* IRES *d2EYFP* mRNA concentration, x_2 the tTA protein concentration, x_3 the unfolded d2EYFP protein concentration and x_4 the folded d2EYFP protein concentration, the PFL network can be described as follows:

$$\frac{dx_1}{dt} = v_1 \left(\alpha_1 + (1-\alpha_1) \frac{\left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} x_2\right)^{h_1}}{K_1^{h_1} + \left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} x_2\right)^{h_1}} \right) - d_1 x_1, \tag{3.1}$$

$$\frac{dx_2}{dt} = v_2 x_1 - d_2 x_2, \tag{3.2}$$

$$\frac{dx_3}{dt} = v_2 x_1 - (d_3 + K_f) x_3, \tag{3.3}$$

$$\frac{dx_4}{dt} = K_f x_3 - d_3 x_4. \tag{3.4}$$

Note that, due to the presence of the IRES sequence, the concentrations of tTA protein and d2EYPF protein depend on the same variable (x_1), that is the concentration of the single mRNA transcript encoding for both proteins.

For the NOPFL network, let x_1 represent only the *d2EYFP* mRNA concentration, and ($\bar{x}_2$) the constant level of the tTA protein, due to the constitutive promoter driving *tTA* expression in the NOPFL cells. The equations thus become:

$$\frac{dx_1}{dt} = v_1 \left(\alpha_1 + (1-\alpha_1) \frac{\left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} \bar{x}_2 \right)^{h_1}}{K_1^{h_1} + \left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} \bar{x}_2 \right)^{h_1}} \right) - d_1 x_1, \quad (3.5)$$

$$\frac{dx_3}{dt} = v_2 x_1 - (d_3 + K_f) x_3, \quad (3.6)$$

$$\frac{dx_4}{dt} = K_f x_3 - d_3 x_4. \quad (3.7)$$

So, the NOPFL is represented by a simpler linear system.

3.2.1 Fitting the model parameters

After deriving the differential equations, an estimation of the equation parameters is needed: in order to obtain a quantitative model of networks, biological experiments are needed to estimate the equation parameters. This can be achieved by minimizing performance measure defined as the error between model predictions and observations, which in this case are the experimental time-series obtained during the "switch off" experiments, depicted in Figure

3.5 and 3.6. Twelve parameters had to be fitted, 11 of which are common to both the PFL and NOPFL models (Table 1).

Parameter were estimated using the Trust Region method (TRM) implemented in PottersWheel [84]; thanks to the multi−model and multi−experimental capabilities of this tool, 12 parameters were identified by simultaneously fitting Eqs. (1) to (7) to all of the experimental time-courses at once. These time-courses include all of the different Doxycycline concentrations for both the PFL and NOPFL cells for a total of 24 time-courses, when taking experimental replicates into account.

Note that the paramter h_2 in Table 1, which affects the strength of Doxycycline repression on the tTA protein activity, is much smaller than 1: usually Hill coefficients are greater than 1, but for the range of Doxycycline concetration used in the experiment (1ng/mL to 1μg/mL), using the parameters' values in Table 1, the function $\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}}$ can be approximated by the function $\alpha + \beta\frac{\theta}{\theta+D}$ ($\alpha \approx 0.4$ and $\beta \approx 0.6$). This means that a Michealis-Menten function can also describe the effect of Doxycyline on tTA activity, but a certain level of leakiness (α) must be taken into account; that is even for large concentrations of Doxycycline, the activity of the tTA protein cannot be completely shut down.

The "switch off" time-series experiments were simulated with both the PFL and NOPFL models using the fitted parameters as shown in Figure 3.5. The inferred models are able to recapitulate the observed dynamics in

response to different inducer concentrations and experimental settings. Note that, since the two networks are constituted by the same biological parts, the parameters for both the PFL and NOPFL models are identical, except for $\bar{x}_2$ in the NOPFL equations, which is not present in the PFL model. Hence, the observed differences in the dynamical behaviour of the PFL and NOPFL networks are due to the intrinsic differences in their topology, and are robust to changes in parameters values.

Experimental determination of the reporter protein degradation

In order to evaluate in a more precise way and from experimental data the degradation rate of the reporter protein (d2EYFP) some experiments were performed on the monoclonal populations. Stably integrated NOPFL cells were treated with Cycloheximide to a final concentration of 10μg/mL, 50μg/mL, 100μg/mL or 500μg/mL, to inhibit protein synthesis [127]. The fluorescence intensity of NOFPL cells was followed for 12 hrs and images were acquired at 15 min intervals. The resulting d2EYFP dynamics are shown in Figure 3.2 and appear very similar, independently of the Cycloheximide concentrations. The experimental data were fitted to an exponential curve $ke^{-d_3 t}$, and the degradation coefficient d_3 was used to obtain the half-life ($\tau_{\frac{1}{2}}$) of the d2EYFP protein: $\tau_{\frac{1}{2}}$=log(2)/ d_3 (Figure 3.2 and Table 1). So, $\tau_{\frac{1}{2}}$ has to be in the range 3.6h-4.4h. The estimated value is about two-fold the reported d2EYFP half-life of 2h [143]; we believe that this discrepancy is likely due to

Table 3.1: Parameters identified after the fitting procedures: parameters values as well as standard deviation are reported for each parameter.

Parameter	Description	Fitted value	STD
K_1 [nM]	Activation coefficient	4.81	1.06
α_1	Basal activity CMV-TET promoter	1.13E-05	3.62E-05
v_1 [nM min^{-1}]	Maximal transcription rate CMV-TET promoter	7.54E-02	1.97E-02
v_2 [min^{-1}]	General translation rate	2.71E-02	1.22E-02
d_1 [min^{-1}]	Degradation rate tTA mRNA	1.01E-02	1.22E-03
d_2 [min^{-1}]	Degradation rate tTA protein	1.00E-02	3.42E-03
d_3 [min^{-1}]	Degradation rate d2EYFP protein	3.24E-03	2.66E-04
h_1	Hill coefficient of the CMV-TET promoter	3.16	1.40E-01
θ [nM]	Affinity Doxycycline CMV-TET promoter interaction	1.00	8.85E-03
K_f [min^{-1}]	Folding rate d2EYFP	1.24E-03	1.41E-02
$\bar{x}_2$ [nM]	Steady state tTA in NOPFL	13.69	7.63E-01
h_2	Hill coefficient for Doxycyline	6.03E-02	7.19E-03

the fact that cells were grown at a temperature 32 °C, rather than the usual 37 °C.

3.3 Dynamical properties of the PFL: mathematical prediction and experimental validation

To observe the dynamics of the PFL and NOPFL networks, time-series experiments were performed in which stably-integrated CHO-PFL cells and CHO-NOPFL cells were imaged using time-lapse fluorescence microscopy. The experimental design consisted in treating both PFL and NOPFL cells with different amounts of Doxycycline in order to "switch off" the circuit, by preventing the tTA protein from binding the *CMV-TET* promoter. The following Doxycyline concentrations were tested: 1ng/mL, 10ng/mL, 100ng/mL and 1μg/mL and the dynamic behaviour of both the PFL and NOPFL cells for 43h was followed, collecting images every 15 min, and quantifying the average fluorescence intensity of the cell population. In this way, we averaged out cell-to-cell variability in the response, since at the beginning of each experiment the tracked microscopy field contained at least 15 cells.

The average fluorescence intensity of the reporter gene across the cell population for both the PFL and NOPFL networks is shown in Figure 3.5 for the different concentrations of Doxycycline indicated. In Figure 3.6 replicate

time-course experiments are shown for each of the Doxycycline concetrations used.

The most striking feature is the slowing down in the switch off time of the PFL cells as compared to the NOPFL cells; moreover, the switch off time of the PFL is affected by Doxycycline concentrations, whereas NOPFL cells always switch off with approximately the same dynamics.

In order to further investigate the relationship between topology and dynamical properties, we first observed that the NOPFL model described by Eq.3.5-3.7 is a system of linear time-invariant ODEs, for which the theory of linear dynamical systems applies [8]. From the theory, we know that changes in Doxycycline concentration in Eq.3.5 will not affect the dynamic behaviour of the model, which is governed by the smallest among three degradation terms $d_1, d_3, (d_3 + K_f)$. The concentration of Doxycycline affects only the steady-state values, i.e. how much the network will switch off, but not its dynamics, i.e. how fast it will switch off. Therefore, independently of the values of the parameters, the model of the NOPFL network predicts that for any concentration of Doxycycline, the network will switch off with the same dynamics, albeit possibly reaching different steady-state levels.

Figure 3.7 reports the "switch off" time, τ_{off}, for both the PFL (dashed) and the NOPFL (solid) networks as a function of Doxycycline concentration, computed via numerical simulations of the two models with the parameters estimated in Table 1. τ_{off} is defined as the time taken by the fluorescence

intensity to reach 50% of its final steady-state value (OFF), following treatment with Doxycycline at a given concentration . As expected, the τ_{off} for the NOPFL network is constant and does not change with Doxycyline. This is in agreement with the experimental observations; in Figure 3.7, the switch off time for the NOPFL network for the different concentration of Doxycycline was estimated from the experimental time-series data (defined by $\times$ in Figure 3.7).

On the other hand, the PFL network has a very different behaviour, as can be seen in Figure 3.7. Specifically, for a range of Doxycycline concentrations, the PFL τ_{off} is considerably longer (defined by $+$ in Figure 3.7) that the NOPFL counterpart, which again is in agreement with the experimentally observed behaviour.

Moreover, in order to investigate the origin of the observed dynamical behavior of the PFL circuit, the nullclines for different Doxycycline concentrations are shown in Figure 3.3. When no Doxycycline is present, two stable points (OFF and ON) and one unstable equilibrium point coexist in the same phase portrait, thus providing evidence for the bistability of the PFL network, a shared property among positive feedback loops [17]. However, as Doxycycline concentration increases, while the bacin of attraction of the OFF state become larger for intermediate value of Doxycycline concentration, for higher value the bistability is lost (Figure 3.3), and the only possible equilibrium point is the OFF state, thus making the system monostable. This

implies that without Doxycycline the probability for a cell to stay in an OFF state is lower than the probability of staying in the ON state but, when cells are cultered with fixed concentration of the external inducer, the probability of the system to stay in an OFF state increases. So, from an experimental point of view is simpler to see two subpopulations for the same clone when cells are threated permanently with Doxycycline: however we know from the mathematical analysis that bistability is also a structural property of the PFL not-depending on the permanent threatment with the external inducer. In order to prove this theoretically and experimentally we first simulate the response of the system when only transient stimulus of different duration of Doxycycline are provided to the system (Figure 3.4, panel b).

So, a simulated *in silico* series of "switch OFF" experiments is provided: the switch OFF simulations were performed starting from the ON steady-state and then simulating treatment with Doxycycline for a limited time interval (Δ).

Numerical simulations of the d2EYFP fluorescence level in Figure 3.4b for the PNFL model, and in Figure 3.4b for the PFL model (used as a control) show that, in both cases, when the duration of Doxycycline treatment is below a threshold value ($\Delta_{th}^{PFL} = 1200$ min), d2EYFP expression initially decreases (switches off) and then increases back ON once Doxycycline is removed. However, when the Doxycycline treatment lasts longer than the

threshold duration Δ_{th}, d2EYFP expression switches OFF and it does not increase again, even after Doxycycline has been removed. This is the telltale feature of a bistable "toggle switch".

To experimentally probe the toggle-switch behaviour of the PNFL circuit, and to confirm numerical simulations, we used an innovative microfluidics platform coupled to an inverted epi-fluorescence microscope to administer Doxycycline and to follow in real time d2EYFP fluorescence in clonal populations of PFL [19], (in particular, on the same clonal population of PFL cells selected before (PFL 7)). We then performed *in vitro* "switch OFF" experiments by treating PFL cells with Doxycycline (at a concentration of $1\mu g/ml$) for a limited time interval (Δ). The duration Δ of Doxycycline treatments was chosen according to the *in silico* analysis performed with the PFL models.

d2EYFP fluorescence measurements (Figure 3.4c,d) confirmed the results of the *in silico* analysis, showing that the PFL is bistable: indeed treatment with a pulse of Doxycycline lasting 1800 min (Figure 3.4d) permanently switched the circuit OFF, whereas a pulse of 960 min is not able to switch off the system completely.

Of course, the dynamics are the same of the other experiments performed without the microfluidic device and with a permanent stimulus of Doxycycline.

3.4 Conclusions

This study demostrates in a mammalian experimental system that a transcriptional positive feedback loop can slow down the "switch off" times, as compared to an equivalent network without auto−regulation.

The reason for a cell to "choose" a PFL control strategy for transcriptional regulation, rather than the NOPFL strategy, could be due to the intrinsic robustness of this approach to transient activation of the network. For example, in a signalling pathway, a ligand (equivalent to Doxycycline in our PFL) could cause a transcription factor to stop transcribing itself, as well as, a set of target genes, to initiate a specific response. However, in order for the pathway not to respond to a transient concentration of the ligand, the PFL strategy has to be chosen, otherwise the response would start immediately (NOPFL case). Moreover, the response time of the PFL network can be modulated by the ligand concentration, if this is really high, the system will switch off as quickly as possible (Figure 3.7), alternatively the ligand can be present at low, or medium, concentration, but it should persist for a long time, in order for the pathway to respond. This kind of behaviour has been recently described as "persistence detection" in cellular signal processing to indicate the ability of the genetic circuit to distinguish between transient and persistent signals [146].

The PFL network can exhibit bistability for zero or low concentrations of Doxycyline (Figure 3.3). A bistable genetic network will cause a population

of cells to divide in two sub-populations, each in one of the two possible states (OFF or ON). In yeast, this has been experimentally verified using a simple PFL based on the rtTA system [17]. In our mammalian PFL, this behaviour has been demonstrated using the microfluidic platform able to transmit a variable stimulus of external inducer to the cells. Without this sort of experiment, it is more difficult to evaluate this behaviour because the basin of attraction of the OFF equilibrium point (Figure 3.3) is much smaller as compared to that of the ON state, when no Doxycycline is present. Therefore, just few cells will be stably in the OFF state and these will not be enough to be significantly detected experimentally. However, for intermediate concentration of Doxycycline (100ng/mL in Figure 3.3) the basin of attraction will be comparable and bistability should be detected experimentally.

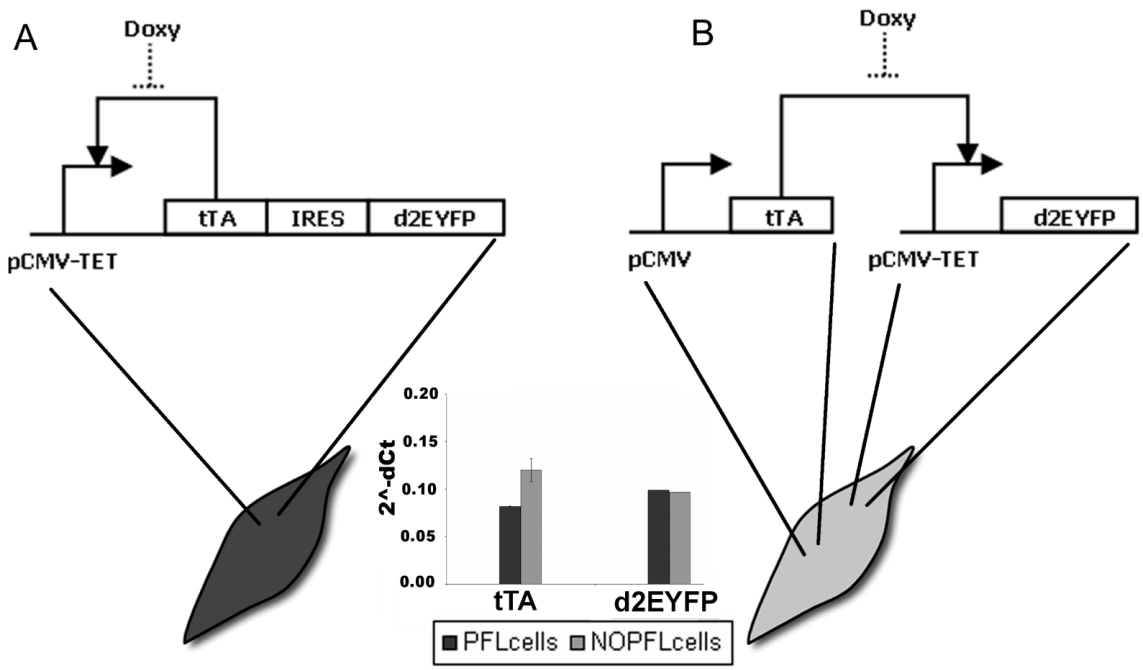


Figure 3.1: **Design of the expression system.** (A) PFL: the promoter *CMV-TET* consists of seven direct repeats of a 42-bp sequence containing the tet operator sequences (tetO), located just upstream of the minimal *CMV* promoter (PminCMV). The Tetracycline-controlled transactivator tTA derives from the addition of the VP16 activation domain to the transcriptional repressor TetR. The d2EYFP is the destabilised yellow-green variant of enhanced green fluorescent protein. (B) NOPFL: the *CMV* promoter drives the expression of the tTA, which in turns drives the transcription of the d2EYFP from the *CMV-TET* promoter. (Inset) RealTime PCR performed on DNA extracted from PFL and NOPFL cells shows that the DNA levels of tTA and d2EYFP are comparable among the two clonal cell populations.

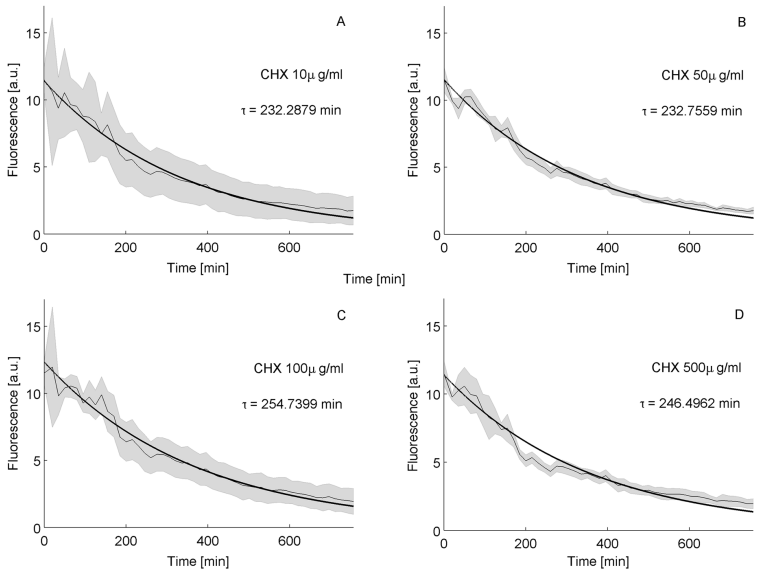


Figure 3.2: **Degradation kinetics of d2EYFP.** *CMV-TET*-d2EYFP stably integrated CHO AA8 TET-OFF cells were treated at t=0 with different concentrations of Cycloheximide (CHX): panel A, 10μg/mL; panel B, 50μg/mL; panel C, 100μg/mL; panel D, 500μg/mL. Fluorescence intensity was followed up to 750 minutes. Sampling time is equal to 15 min. The thin line represents the mean over biological triplicates; the shaded area represents the standard error. Experimental data were used to fit the exponential decay of d2EYFP protein levels, and thus to derive its half-life (τ). Fluorescence intensity in untreated cells was not subjected to any significant decay (data not shown).

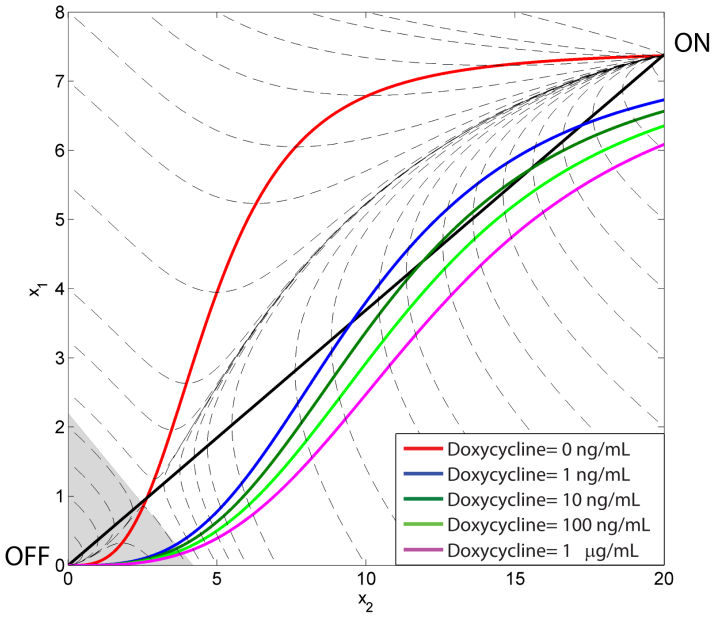


Figure 3.3: **Phase portrait of the PFL model.**The tTA-d2EYFP mRNA concentration (y axis) has been plotted against tTA protein concentration (x axis). Varying Doxycycline concentrations (1 ng/mL through 1 μg/mL) were used to investigate the dependence of the two stable equilibria ("ON" and "OFF" in the graph) on the amplitude of the input. The shape and dimensions of the two basins of attraction (the set of initial conditions ending up in one of the two stable steady states) can be studied with the same technique: in this figure the grey shaded area represents the basin of attraction of the "OFF" equilibrium for Doxycycline= 0 nM.

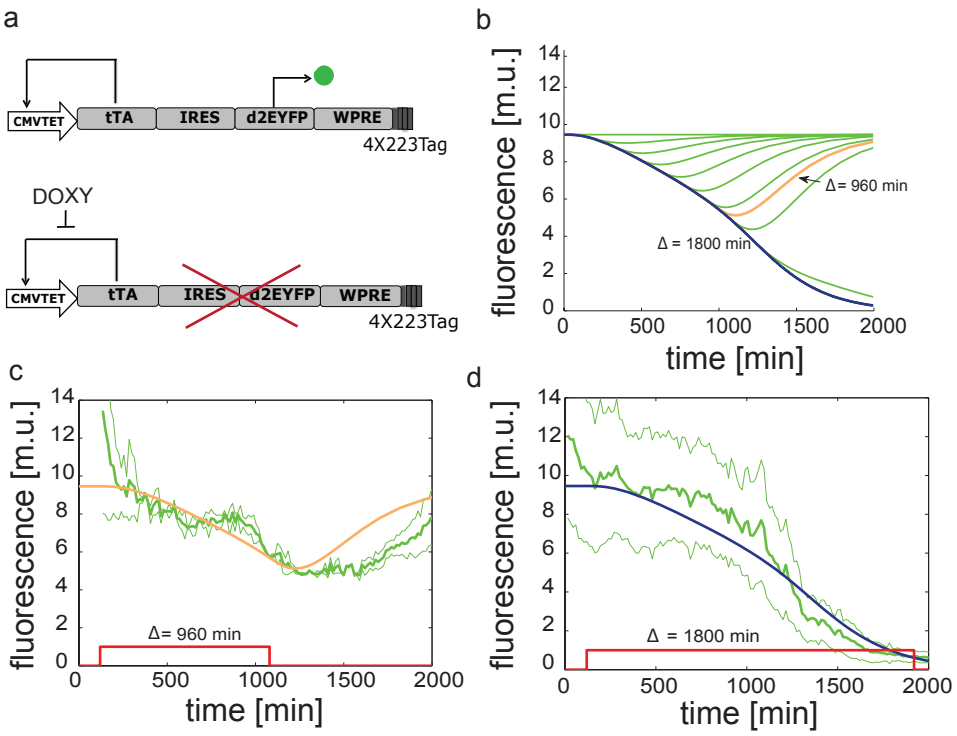


Figure 3.4: **In silico and *in vitro* "switch OFF" experiments for PFL cells following pulses of Doxycycline of different duration. (a)**: The positive feedback loop (PFL) exogenously controlled by Doxycycline able to switch off the system. **(b)**: Simulated d2EYFP fluorescence of PFL cells following simulated treatment with Doxycycline of different duration. **(c)** and **(d)**: experimental d2EYFP fluorescence using the microfluidics device (solid green line) following treatment with Doxycycline (red line) at time 120 min and removed after $\Delta = 960$ (c) or 1800 min (d); standard deviation (thin green lines) is among at least three replicates; simulations (blue and orange lines) are rescaled to experimental data and also represented in (b) (same colors).

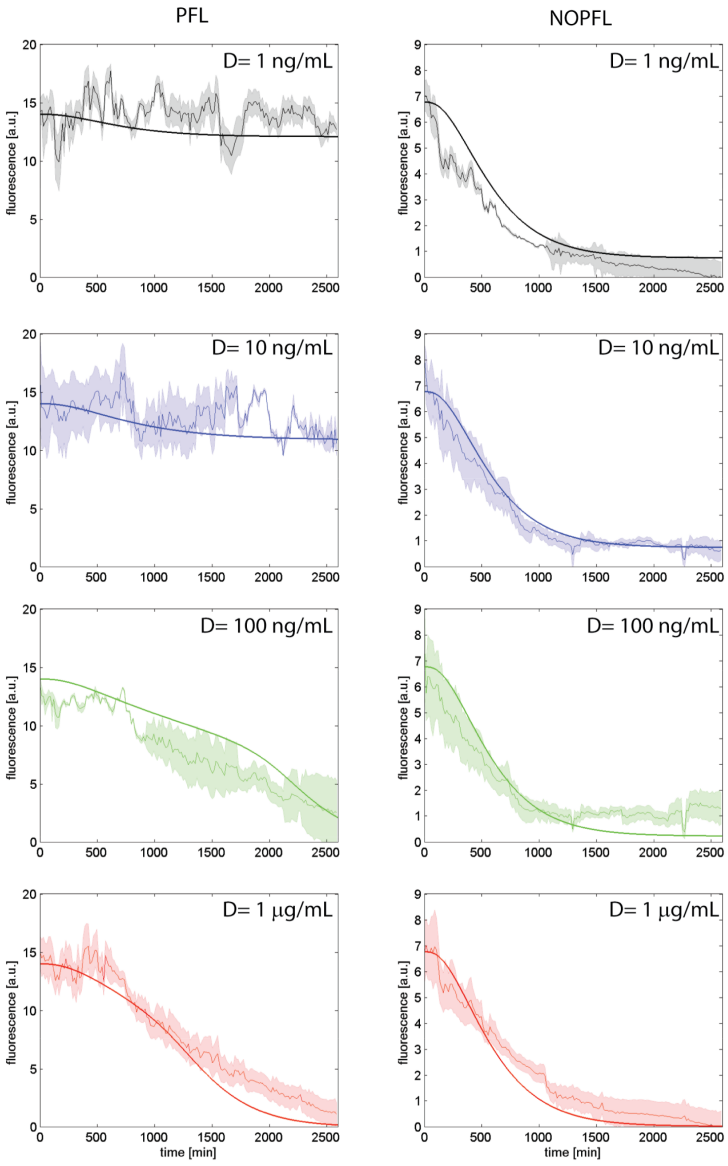


Figure 3.5: **Experimental and simulated switch off time-course across the PFL and NPFL cell population.** Experimental data (thin lines) and model simulations (thick lines) were reported for the PFL (left) and NOPFL (right) cells. Shaded areas represent standard deviations from replicate experiments.

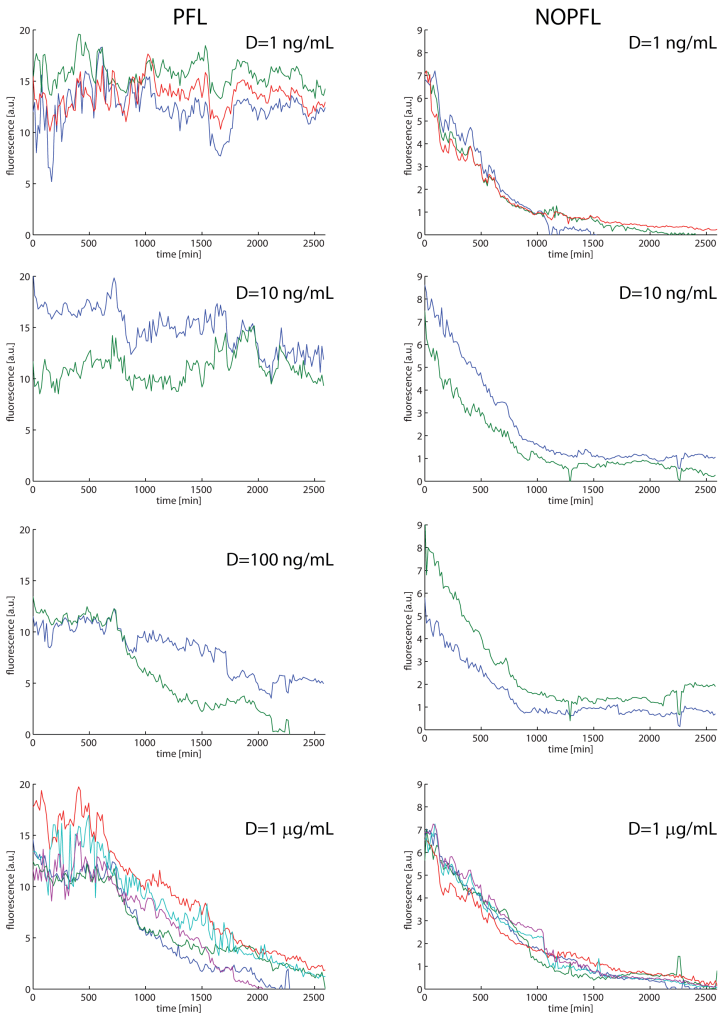


Figure 3.6: **Replicates of the experimental time-courses across the PFL and NPFL cell population.** Replicates of the experimental time-courses for the PFL (left) and NOPFL (right) cells. Each line in each panel represent the average fluorescence intensity across the cell population in one switch-off experiment.

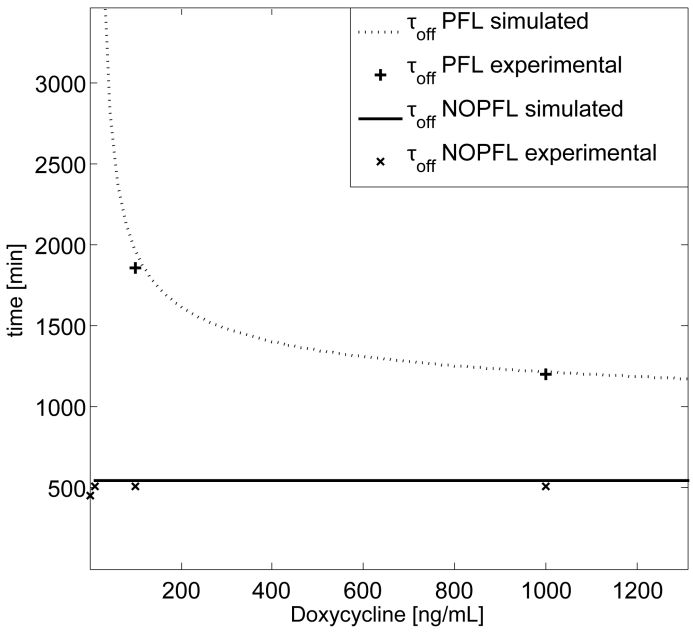


Figure 3.7: **Switch off time τ_{off} for varying Doxycycline concentrations from experimental data and model predictions.**The model predictions for the switch off times τ_{off} are shown for PFL (dashed thick line) and NOPFL (solid line). Experimental quantification of the τ_{off} for PFL and NOPFL models have been reported for comparison with + and $\times$ respectively. Observe that the experimental τ_{off} for the PFL at 1ng/mL and 10ng/mL could not be estimated since the PFL is not switching off in the experimental observation time (43h).

Chapter 4

Modeling of a Positive and microRNA-based Negative Feedback Loop system

4.1 Introduction

Negative feedbacks play a central role in trascriptional and post-trascriptional regulation: autoregulatory, negative feedback loops in gene circuits provide stability, thereby limiting the range over which the concentrations of network components fluctuate ([6], [60], [16]). Negative feedbacks can have different roles: delayed negative feedbacks are known to generate oscillations [99], [59], [57] otherwise they can buffer noise fluctuations and also speed up the dynamics of a switch [117].

In Chapter 3, the role of a positive feedback loop has been discussed through mathematical modeling and experimental validation of a simple synthetic circuit inserted in a mammalian cell line. Now, a study is presented in which the same system is considered with the addition of a negative feedback loop mediated by a microRNA. In particular, the construction and the characterization of two synthetic networks composed by the PFL (previously described) combined respectively with a non delayed and delayed microRNA-based Negative Feedback Loop, are here described. The experimental construction of these networks has been performed in the Systems and Synthetic Biology Lab of Dr. Diego di Bernardo by Dr. Velia Siciliano (for further information see http://dibernardo.tigem.it). Both systems have been inserted in a mammalian cell-line. The mathematical characterization of these two constructs reveals that both circuits can present a monostable, bistable or oscillatory behaviour depending on the choice of the parameters: this implies that depending on the physical characteristics of the components of the constructs is possible to obtain a switch (parameters in the bistability region) or a genetic oscillator (parameters taken in the oscillatory region).

4.2 Design of a Positive and Negative Feedback Loop system in mammalian cells (PNFL)

The first network taken into account is the circuit shown in Figure 4.1. This circuit is composed by an activator, which promotes its own transcription, as well as the transcription of a microRNA directed against itself: the activator is the tTA protein, also present in the PFL network, already described in Chapter 3; the same transcription factor also binds another CMV-TET promoter driving the expression of a natural occuring mammalian microRNA (miR-223). Also in this case, the binding of the transcription factor to the promoter can be inihibited by Doxycycline. miR-223 is specifically expressed in the myeloid lineage, where it targets multiple genes inducing a low magnitude of repression, mainly acting as a rheostat to adjust protein output [13]. The microRNA, once expressed, targets the mRNA encoding for the transcription factor, thanks to four sequence tags at the 3' end of the transcript perfectly complementary to the miR223 seed sequence [25].

The PNFL motif, made up of well known and characterized biological parts, orthogonal to the endogenous ones was inserted in Chinese Hamster Ovary (CHO) cells: the positive feedback motif and negative feedback motif are embedded in two separate lentiviral vectors. Also in this case, using the lentiviruses and starting from the 9 PFL monoclonal populations generated as explained in Chapter 3, it was possible to generate fourteen matched PNFL

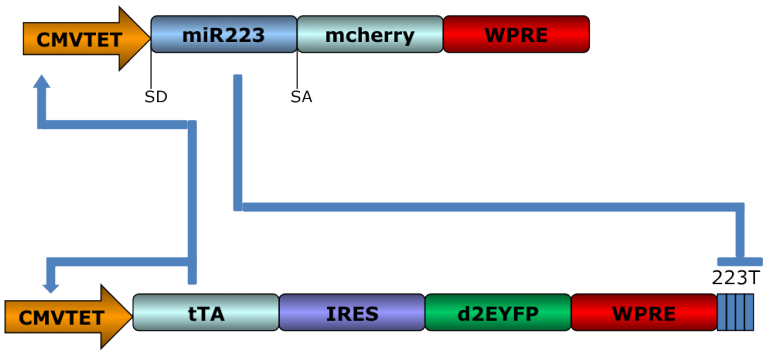


Figure 4.1: **Schematics of PNFL expression system.** : the Tetracycline-controlled transactivator (tTA) is self-regulated by binding the tTA-responsive *CMV-TET* promoter, thus generating a positive feedback loop (PFL - black lines), whose dynamics are tracked by a destabilized EYFP (d2EYFP). The same *CMV-TET* promoter drives the transcription of the human microRNA miR-223 embedded in the first intron of the low affinity nerve growth factor receptor (ΔLNGFR) [25], followed by a reporter gene encoding for the mCherry fluorescent protein (Negative Feedback Loop - NFL - red lines). miR223 in turn down-regulates the tTA mRNA levels through 4-repeated target sequences perfectly complementary to the miR-223 seed sequence, placed at the 3'UTR of the PFL gene expression cassette, thus inducing degradation of the target mRNA. Addition of doxycycline interrupts the tTA-mediate activation of PNFL. WPRE, woodchuck hepatitis virus post-transcription regulatory element.

clonal populations of CHO cells (PNFL cells), carrying both the PFL and the NFL cassettes, as shown in Figs.4.13a and 5.5.

4.2.1 Derivation of the mathematical model

The PNFL network can be represented by different models. In this case I considered an ODE model consisting of a system of seven ODEs as described below

$$\begin{aligned}
\frac{dx_1}{dt} &= G_1 v_1 \left(\alpha_1 + (1-\alpha_1) \frac{\left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}} x_2\right)^{h_2}}{K_1^{h_2} + \left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}} x_2\right)^{h_2}} \right) - d_1 x_1 \qquad (4.1)\\
&- \lambda \frac{{x_4}^{h_3}}{{K_3}^{h_3} + {x_4}^{h_3}} x_1,\\
\frac{dx_2}{dt} &= v_2 x_1 - d_2 x_2, \qquad (4.2)\\
\frac{dx_3}{dt} &= G_2 v_1 \left(\alpha_1 + (1-\alpha_1) \frac{\left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}} x_2\right)^{h_2}}{K_1^{h_2} + \left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}} x_2\right)^{h_2}} \right) \qquad (4.3)\\
&- (d_3 + K_D) x_3,\\
\frac{dx_4}{dt} &= K_D x_1 - \delta x_4, \qquad (4.4)\\
\frac{dx_5}{dt} &= v_4 x_1 - (K_{fg} + d_4) x_5, \qquad (4.5)\\
\frac{dx_6}{dt} &= K_{fg} x_5 - d_4 x_6, \qquad (4.6)\\
\frac{dx_7}{dt} &= v_3 x_3 - (K_{fr} + d_5) x_7, \qquad (4.7)\\
\frac{dx_8}{dt} &= K_{fr} x_7 - d_5 x_8. \qquad (4.8)
\end{aligned}$$

where

- x_1 refers to tTA mRNA,
- x_2 refers to tTA protein,
- x_3 refers to miR223 mRNA,
- x_4 refers to miR223 mRNA in the active form
- x_5 refers to d2EYFP unfolded protein,
- x_6 refers to d2EYFP folded protein,
- x_7 refers to mcherry unfolded protein,
- x_8 refers to mcherry folded protein.

In this model, I considered two different equations for the tTA (one for the mRNA and one for the mature protein), two equations for the microRNA and two different equations for each fluorescent reporter. Indeed, I assumed that the microRNA maturation time can be modeled by an extra linear equation, as well as the fluorescent protein folding.

The structure of this model is the same as the PFL described in Chapter 3 except for the inclusion of an additional term representing the post-transcriptional degradation mediated by the microRNA (x_3):

$$-\lambda \frac{{x_3}^{h_3}}{{K_3}^{h_3} + {x_3}^{h_3}} x_1.$$

The parameter λ represents the strength of the microRNA-induced degradation: when $\lambda = 0$ the PNFL model converges to the PFL one. Another difference with the PFL model, are the additional parameters G_1, G_2: I took into account that the PFL cassette and the Negative Feedback cassette can be inserted into different positions in the genome for the different clonal population. Depending on the insertion site, the genes part of a network can be more or less expressed. Hence by modulating these two parameters it is possible to account for the resulting differences in gene expression across clonal populations.

As a second model, I propose a set of DDE equation, to model this maturation time. In a DDE model in which the delay is explicit and I added it to the tTA mRNA equation.

$$
\begin{aligned}
\frac{dx_1}{dt} &= G_1 v_1 \left(\alpha_1 + (1-\alpha_1) \frac{(x_2)^{h_2}}{K_1^{h_2} + (x_2)^{h_2}} \right) - d_1 x_1 && (4.9)\\
&- \lambda \frac{x_3(t-\tau)^{h_3}}{K_3{}^{h_3} + x_3(t-\tau)^{h_3}} x_1, \\
\frac{dx_2}{dt} &= v_2 x_1 - d_2 x_2, && (4.10)\\
\frac{dx_3}{dt} &= G_2 v_1 \left(\alpha_1 + (1-\alpha_1) \frac{(x_2)^{h_2}}{K_1^{h_2} + (x_2)^{h_2}} \right) - d_3 x_3, && (4.11)\\
\frac{dx_4}{dt} &= v_4 x_1 - (K_{fg} + d_4) x_4, && (4.12)\\
\frac{dx_5}{dt} &= K_{fg} x_4 - d_4 x_5, && (4.13)\\
\frac{dx_6}{dt} &= v_3 x_3 - (K_{fr} + d_5) x_6, && (4.14)\\
\frac{dx_7}{dt} &= K_{fr} x_6 - d_5 x_7. && (4.15)
\end{aligned}
$$

where

- x_1 is the tTA mRNA,
- x_2 is the tTA protein,
- x_3 is the miR223 mRNA,
- x_4 is the d2EYFP unfolded protein,
- x_5 is the d2EYFP folded protein,
- x_6 is the mcherry unfolded protein,
- x_7 is the mcherry folded protein.

Also in this case the model is structured in the same way as before, but the delay is explicit in the action of the microRNA degradation.

4.2.2 Bifurcation analysis and results

It is possible to perform a bifurcation analysis in order to derive what differences in the properties come from the difference in the modelling approach. The considered models have a large number of parameters common to the PFL model (i.e. the production and degradation rate of the tTA and d2EYFP and the Hill functions describing the tTA self-activation and the effect of Doxycycline). For the choice of the parameters concerning the microRNA, it is useful to start with parameters from literature: in [34], a mathematical model is presented to simulate the effect of silencing RNAs on gene expression. The first choice of parameters I used to simulate the models is reported in Table 4.1: with these parameters the models do not present an oscillatory behaviour, on the contrary they exhibit two different stable equilibria (so they are bistable). In Figure 4.2 the bifurcation diagram for the DDE model is reported; the ODE model present a similar bistable behaviour as shown in the inset (data not shown).

I observed that slightly moving some parameters in the DDE model from their initial values it is possible to obtain a stable focus in the phase plane starting from the bistability region (Figure 4.2). In particular, the parameters to modify are those concerning the folding and the action of the microRNA:

$d_3 = 0.108$, $\lambda = 5$, $K_3 = 0.007$, $K_D = 0.00009$, $h_3 = 7$, $\alpha = 0.085$. With this choice of parameters only the DDE model exhibit damped oscillations, but the same is not true for the ODE model. Moving h_3 to higher values it is possible to reach a limit cycle via a Hopf bifurcation, as shown in the continuation diagram shown in Figure 4.2.

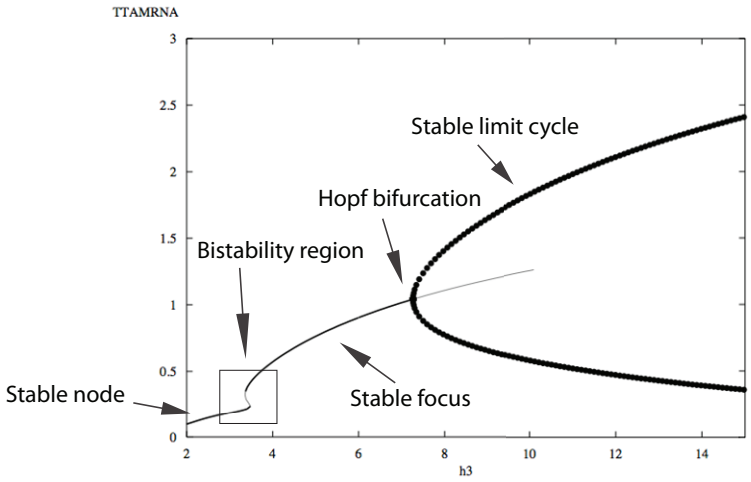


Figure 4.2: **Bifurcation diagram for DDE model with parameters of Table 4.2.** tTA mRNA steady state is plotted against the parameter h_3 thus revealing a possible bistable behaviour for low values of the parameter and the existance of a limit cycle for higher values.

I then performed a 2-parameter bifurcation analysis by constraining the h_3 parameter to be in the limit cycle region in order to get biologically plausible values for the parameters: bifurcation diagrams are shown in Figs. 4.4, 4.5, 4.6, 4.8, 4.9, 4.10, while simulations I obtained with the new set of parameters,

reported in Table 4.2, are shown in Figure 4.3.

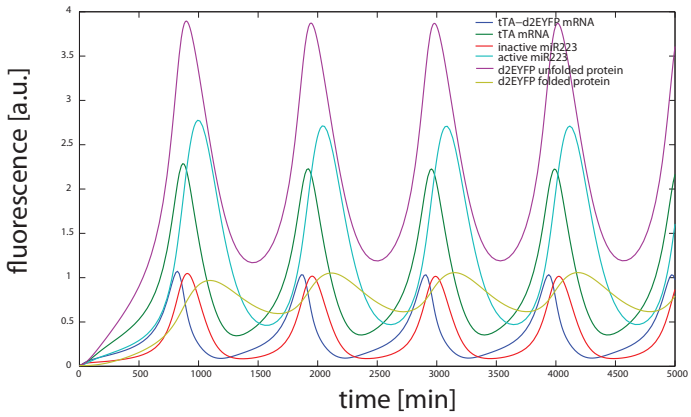


Figure 4.3: **Simulation of the DDE model in the oscillatory region.** Simulation over time of the oscillatory behaviour of the PNFL motif when parameters are chosen in the region in which a limit cycle exists (Table 4.2).

For this new choice of parameters (Table 4.2), I also simulated the behaviour of the ODE model but it does not present any oscillation. Althought I cannot conclude that the ODE model is not able to present an oscillatory behaviour, I can conclude that the structure of the DDE system is able to exhibit bistability or limit cycle oscillations. In particular, the values reported in Table 4.2 are compatible with biological constraints and the convergence to a limit cycle or to a fixed point is robust enough to perturbations as the bifurcation diagrams show (Figs 4.4, 4.5, 4.6, 4.8, 4.9, 4.10 - larger markers correspond to longer periods).

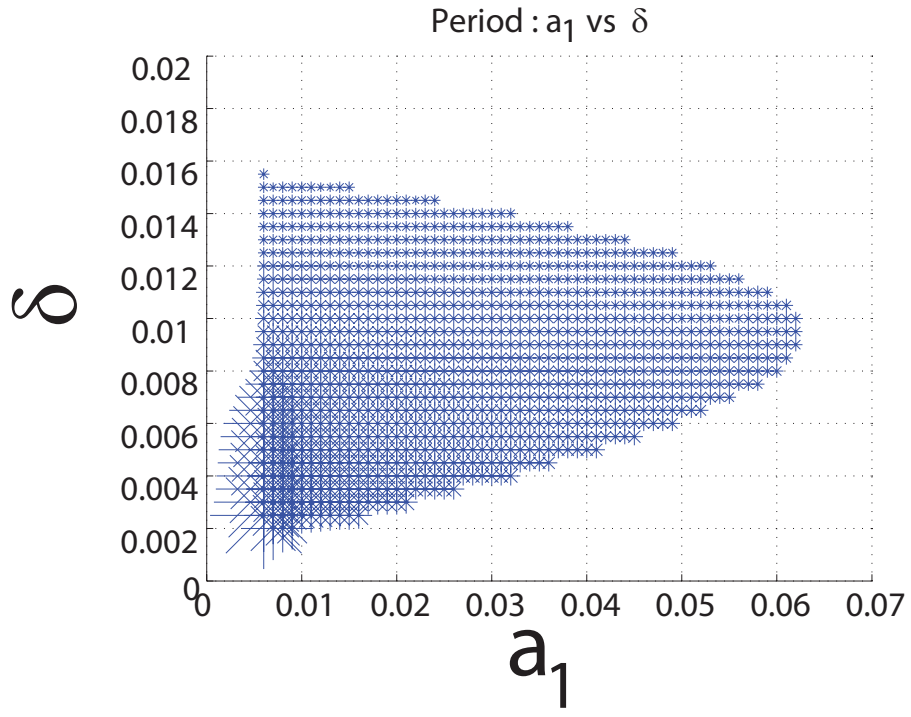


Figure 4.4: **Bifurcation diagram for α_1 and δ, DDE model.** Two variables bifurcation diagram between the leakiness of the tTA promoter and the degradation of the microRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

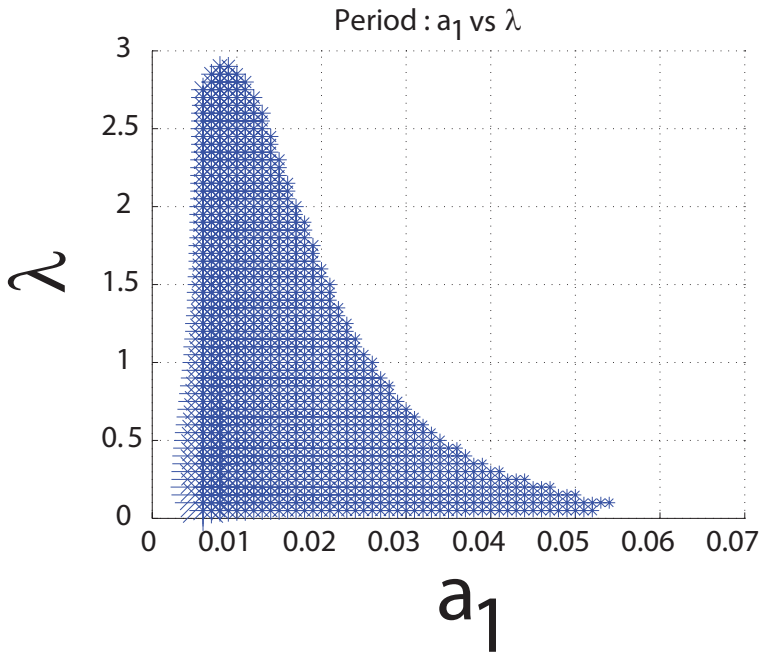


Figure 4.5: **Bifurcation diagram for α_1 and λ, DDE model.** Two variables bifurcation diagram between the leakiness of the tTA promoter and the strength of post-trascriptional degradation of the microRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

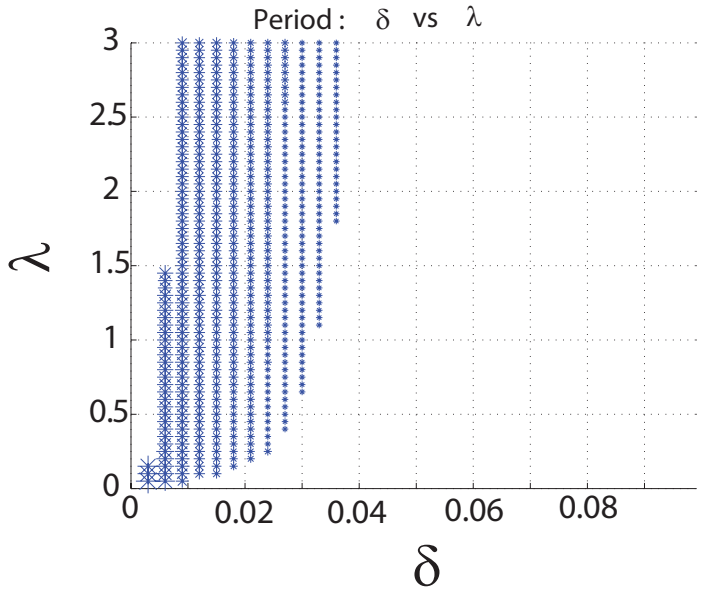


Figure 4.6: **Bifurcation diagram for δ and λ, DDE model.** Two variables bifurcation diagram between the degradation and the strength of post-trascriptional degradation of the microRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

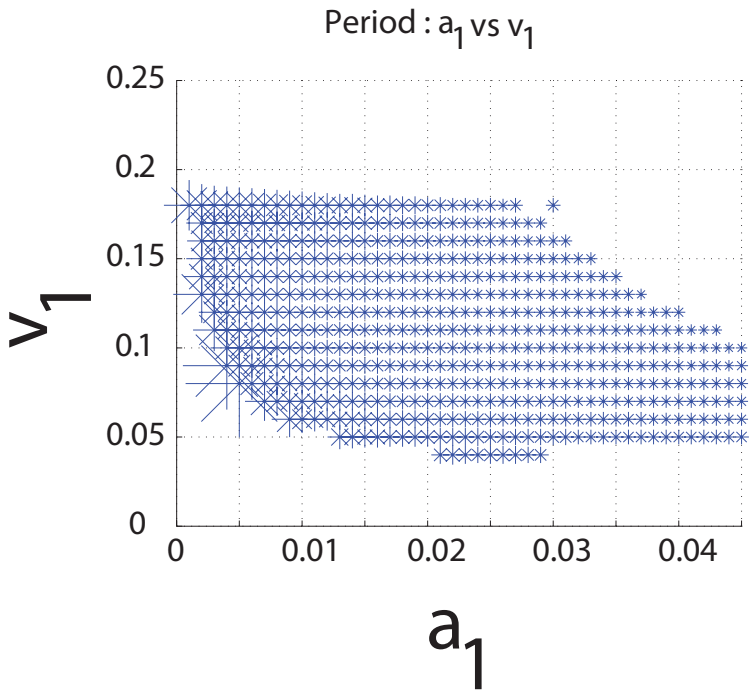


Figure 4.7: **Bifurcation diagram for a_1 and v_1, DDE model.** Two variables bifurcation diagram between the leakiness and te maximal transcription rate of the tTA promoter: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

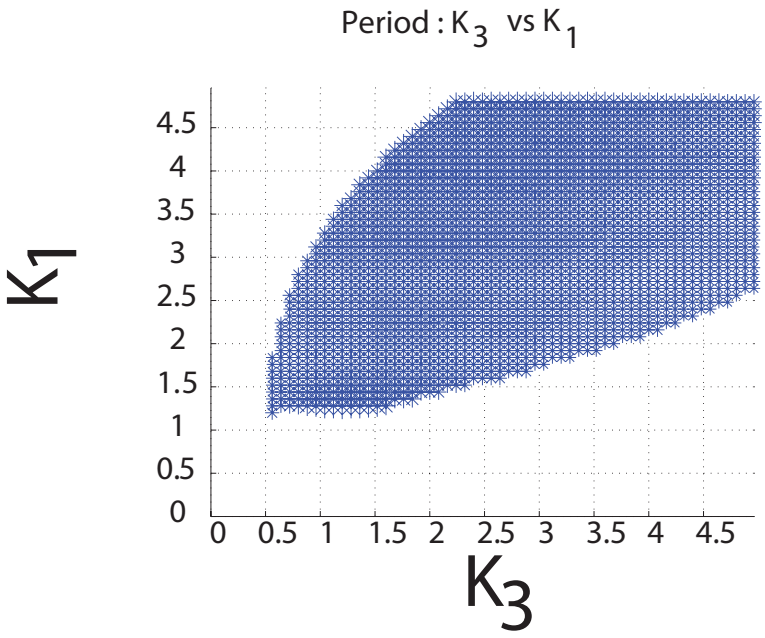


Figure 4.8: **Bifurcation diagram for K_3 and K_1, DDE model.** Two variables bifurcation diagram between the Michelis-Menten constants of the tTA and the miRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

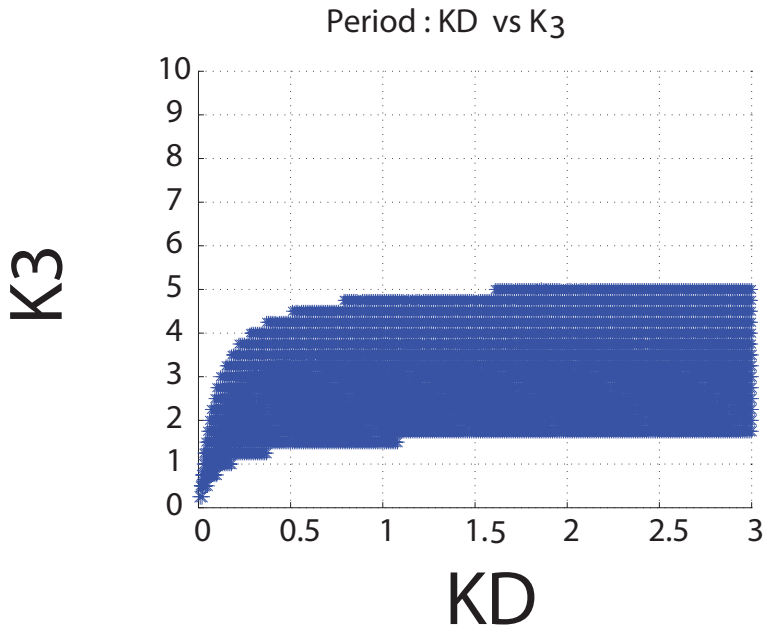


Figure 4.9: **Bifurcation diagram for K_3 and K_D, DDE model.** Two variables bifurcation diagram between the Michelis-Menten constant and the maturation rate of the microRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

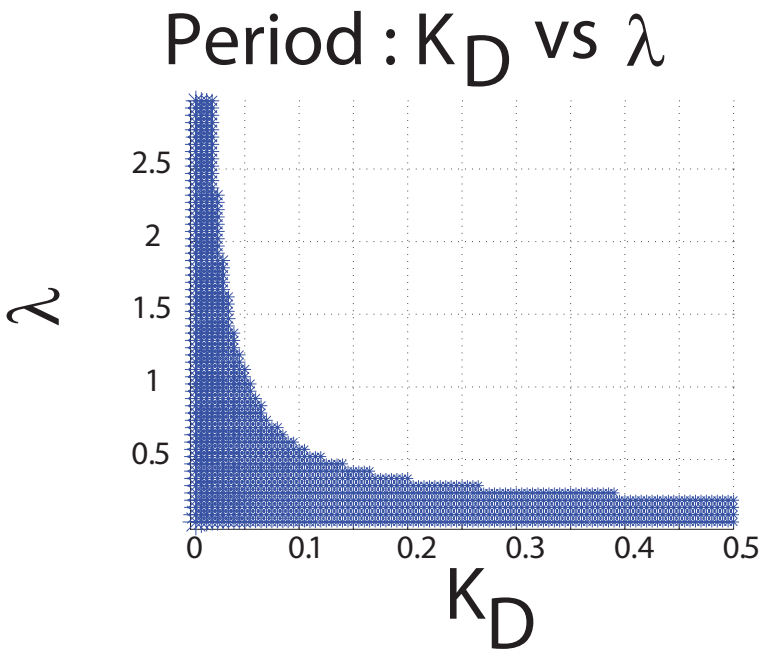


Figure 4.10: **Bifurcation diagram for K_D and λ, DDE model.** Two variables bifurcation diagram between the Michelis-Menten constant and the post-transcriptional degradation strength of the microRNA: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

4.3 PNFL as a toggle-switch: experimental validation

In parallel with the construction of a mathematical model, the generation of 14 monoclonal populations carrying the PNFL construct was performed. The screening of all these clones by time-lapse experiments revealed that most of them does not exhibit any oscillatory behaviour, see [124]. Hence, clones were analysed to study if some of them did present a bistable behaviour. In particular, a switch off time-lapse experiment was performed on clone PNFL 7-2 so to set the PNFL parameters in a compatible manner with the clone considered. Thus, the parameters not belonging to the PFL were chosen to minimize the Least Squared Error of the d2EYFP fluorescence time-course following Doxycyline administration to the cell, as reported in Figure 4.15.

A bifurcation analysis shows that, with these parameters, the ODE model presents a bistable behaviour as shown in 4.14; this is true also for DDE model 4.12. However, it is convenient to choose equations 4.1 to describe the system, first thanks to the presence of only ordinary differential equations (simpler for mathematical studies) and second because it is known, from a biological point of view, that, effectively, microRNA maturation takes more steps ([73], [28], [25], [35], [68]).

In order to detect in vivo this predicted bistable behaviour, a series of "switch off" experiments was simulated for both the PFL and PNFL motifs:

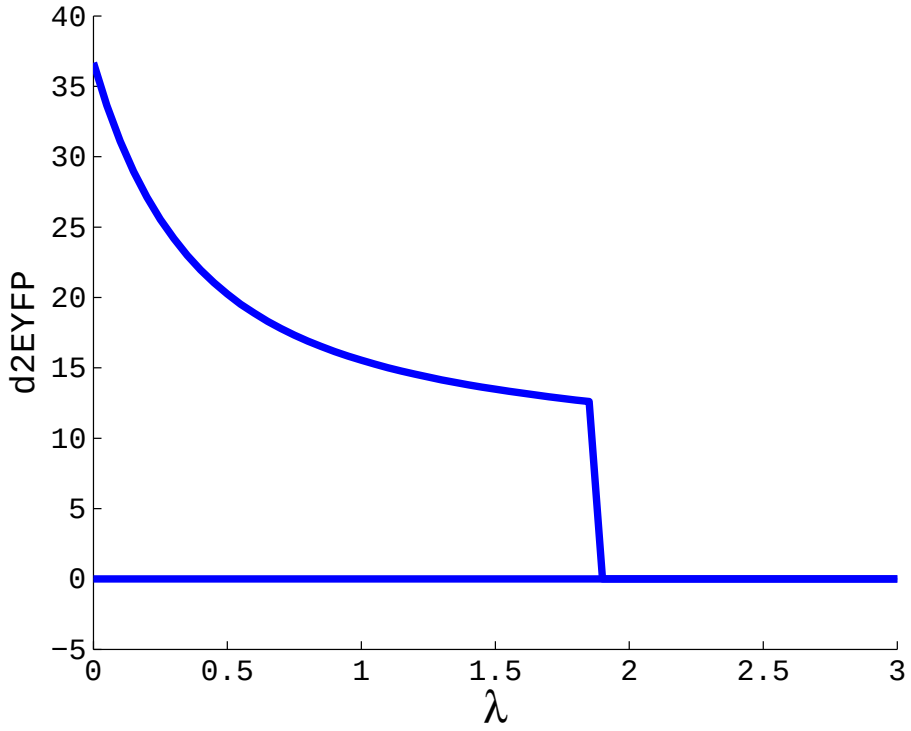


Figure 4.11: **Bifurcation diagram for the PNFL described by the first ODE system.** The tTA mRNA is plotted as dependent on the microRNA strength as described in the first model. The system reveals a bistable behaviour for low values of the parameter (λ).

the switch off simulations were performed by setting as initial conditions for solving Eqs. 4.1-4.1, the values of the state variables at the ON steady state, we then simulated pulses of Doxycycline treatment (at a concentration of $1\mu g/ml$) of different durations Δ.

The behaviour of the PNFL motif and its robustness to model parame-

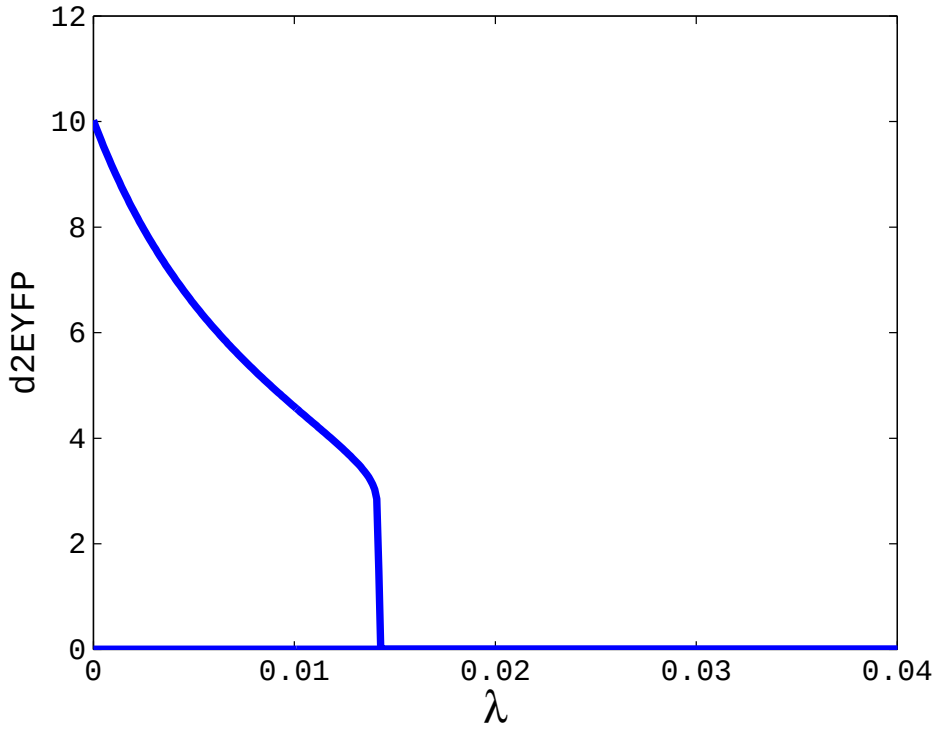


Figure 4.12: **Bifurcation diagram for the PNFL described by the DDE model.** The tTA mRNA is plotted as dependent on the microRNA strength as described in the second model. The system reveals a bistable behaviour for low values of the parameter (λ).

ters can be explored by numerical bifurcation analysis. As shown in the inset of Figure 4.13b the PNFL can behave as a bistable system (toggle-switch), which, in the absence of Doxycycline, can either be ON (high expression of the d2EYFP) or OFF (low expression of the d2EYFP), as long as the repression of the miRNA is not too strong. Specifically, when the level of the TF

is above the promoter activation threshold, the PFL motif will switch ON and stay ON due to its autocatalytic activity; if, however, the TF activity is transiently blocked (i.e. by a small molecule, or signaling event) then its protein product will decrease below the promoter activation threshold and the PFL will switch OFF and stay OFF even after the transient block has been removed [147], [51], [83], [17], [126], [78].

Numerical simulations of the d2EYFP fluorescence level in Figure 4.13b for the PNFL model, and in Figure 4.13b for the PFL model (used as a control) show that, in both cases, when the duration of Doxycycline treatment is below a threshold value ($\Delta_{th}^{PNFL} = 50$ min and $\Delta_{th}^{PFL} = 1200$ min), d2EYFP expression initially decreases (switches off) and then increases back ON once Doxycycline is removed. However, when the Doxycycline treatment lasts longer than the threshold duration Δ_{th}, d2EYFP expression switches OFF and it does not increase again, even after Doxycycline has been removed. This is the telltale feature of a bistable "toggle switch".

To experimentally probe the toggle-switch behaviour of the PNFL circuit, and to confirm numerical simulations, a microfluidics platform was used, coupled to an inverted epi-fluorescence microscope to administer Doxycycline and to follow in real time d2EYFP fluorescence in clonal populations of PFL and PNFL cells [19]. PFL7 clone and the matched clonal population of PNFL cells (PNFL7-2) derived from the same PFL clone were used to monitor the

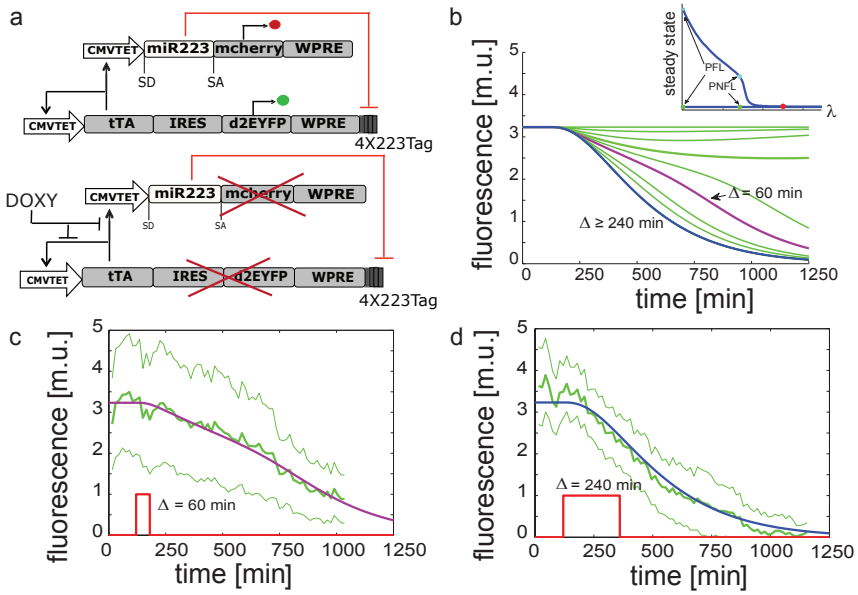


Figure 4.13: **In silico and *in vitro* "switch OFF" experiments for PNFL cells following pulses of Doxycycline of different duration.** **(a)**: Transcription Factor (TF) self-activating its own transcription and the transcription of a microRNA directed against itself; **(b)**: Simulated d2EYFP fluorescence of PNFL cells following simulated treatment with Doxycycline of different duration Δ. **(c)** and **(d)**:experimental d2EYFP fluorescence using the microfluidics device (solid green line) following treatment with Doxycycline (red line) at time 120 min and removed after $\Delta = 60$ min (c) or 240 min (d); standard deviation is among at least three replicates (thin green lines); simulations (blue and purple lines) are rescaled to the experimental data and also represented in (b) (same colors).

fluorescence. We then performed *in vitro* "switch OFF" experiments by treating PNFL and PFL cells with Doxycycline (at a concentration of $1\mu g/ml$) for a limited time interval (Δ). The duration Δ of Doxycycline treatments was chosen according to the *in silico* analysis performed with the PNFL and PFL models (Figure 4.13 b). Simulation are measured in Model Units (m.u.) and experiments in Arbitrary Units (a.u.). The simulated data were rescaled to the experimental units by using a linear function ($y_{rescaleddata} = ax_{data} + b$). For PFL, the rescaling factors a and b are $(3, -5)$ and $(1.847, 0)$ respectively for $\Delta = 960$ min and $\Delta = 1800$ min; for PNFL, the rescaling factors are $(2.15, 2.82)$ and $(1.92, 1.32)$ respectively for $\Delta = 240$ min and $\Delta = 60$ min.

Numerical simulations were run using Matlab 2010b (Mathworks Inc.), $ode23s$ solver (a detailed discussion of the numerical methods used by $ode23$ can be found in [22]). For the parameter identification, also the PottersWheel toolbox [85] implemented in MATLAB is used, as also described in (11.1).

d2EYFP fluorescence measurements (Figure 4.13 c,d for the PNFL circuit, and Figure 3.4 c,d for the PFL) confirmed the results of the *in silico* analysis, showing that the PNFL is bistable: indeed treatment with a pulse of Doxycycline lasting 60 (Figure 4.13c), or 240 (Figure 4.13d) minutes permanently switched the circuit OFF, whereas a pulse of 20 minutes did not cause any visible effect (data not shown). The PFL had the same overall behaviour but its dynamics were much slower (Figure 3.4c,d). Moreover, as detailed in

the next section, we further probed the bistable behaviour of additional PFL and PNFL clones by Fluorescent Activated Cell Sorter (FACS), as reported in Figure 5.1.

4.4 Design of a Positive and Delayed amplified Negative Feedback Loop in mammalian cells (PDNFL)

In this section I derived for a model for a new design of the Negative Feedback Loop obtained by introducing in the feedback loop another "piece" to "slow down" the negative feedback thus favouring oscillations rather than bistability.

As shown in Figure 4.16, a three stage genetic oscillator was designed to increase the time elapsing between the auto-activation of the positive feedback loop and the repression of the microRNA on the tTA.
The PFL is conserved, but the tTA now drives the expression of a gene expression cassette consisting of:

- destabilized artificial activator (dGAL4-VP16) whose expression has been optimized for mammalian cells [142];
- IRES sequence;

- destabilized cyan reporter (dCyan).

The NFL was modified so that the miR223 is now transcribed from the GAL4 responsive promoter (UAS). This promoter is composed by a minimal cytomegalovirus promoter (CMVmini) fused with four tandem repeats of the galactose upstream activating sequence [142]. Once the dGAL4-VP16 is produced, it binds the responsive promoter (UAS) driving the expression of miR223, thus giving rise to a new Negative Feedback Loop (UAS-NFL), that will silence the tTA-d2EYFP mRNA expression.

4.4.1 Derivation of the mathematical model of the three steps oscillator

I derived a mathematical model based on ODEs. It preserves the structure of the PFL and the part characterizing the cassette containing the microRNA, as derived in Section 4.3.

$$
\begin{aligned}
\frac{dx_1}{dt} &= v_1\left(\alpha_1 + (1-\alpha_1)\frac{\left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_2\right)^{h_2}}{K_1^{h_2} + \left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_2\right)^{h_2}}\right) - d_1x_1 \\
&- \lambda\frac{{x_5}^{h_3}}{{K_3}^{h_3} + {x_5}^{h_3}}x_1, \qquad (4.16)\\
\frac{dx_2}{dt} &= v_2x_1 - d_2x_2, \qquad (4.17)\\
\frac{dx_3}{dt} &= v_1\left(\alpha_1 + (1-\alpha_1)\frac{\left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_2\right)^{h_2}}{K_1^{h_2} + \left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_2\right)^{h_2}}\right) \\
&- d_Gx_3, \qquad (4.18)\\
\frac{dx_4}{dt} &= v_{PG}x_3 - d_{PG}x_4, \qquad (4.19)\\
\frac{dx_5}{dt} &= v_{UAS}\left(\alpha_{UAS} + (1-\alpha_{UAS})\frac{\left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_4\right)^{h_{UAS}}}{K_{UAS}^{h_{UAS}} + \left(\frac{\theta^{h_0}}{\theta^{h_0}+D^{h_0}}x_4\right)^{h_{UAS}}}\right) \\
&- d_3x_5, \qquad (4.20)\\
\frac{dx_6}{dt} &= v_4x_1 - (K_{fg}+d_4)x_6, \qquad (4.21)\\
\frac{dx_7}{dt} &= K_{fg}x_6 - d_4x_7, \qquad (4.22)\\
\frac{dx_8}{dt} &= v_3x_5 - (K_{fr}+d_5)x_8, \qquad (4.23)\\
\frac{dx_9}{dt} &= K_{fr}x_8 - d_5x_9. \qquad (4.24)\\
\frac{dx_{10}}{dt} &= v_5x_3 - (K_{cr}+d_6)x_{10}, \qquad (4.25)\\
\frac{dx_{11}}{dt} &= K_{cr}x_{10} - d_6x_{11}. \qquad (4.26)
\end{aligned}
$$

where

- x_1 is the tTA mRNA,

- x_2 is the tTA protein,
- x_3 is the dGal4 mRNA,
- x_4 is the dGal4 protein,
- x_5 is the miR223,
- x_6 is the d2EYFP unfolded protein,
- x_7 is the d2EYFP folded protein,
- x_8 is the mcherry unfolded protein,
- x_9 is the mcherry folded protein,
- x_{10} is the cyan unfolded protein,
- x_{11} is the cyan folded protein,

This model only contains only one equation for the miR223; the second "motivation" equation is not needed, being the delay produced by the new "pair". A phase-space portrait for this model is reported in Figure 4.17. Unknown parameters were chosen from literature, in particular the half-life of the protein dGAL4 is known to be of 3.76 hours, while the half-life of miR223 was estimated to be of 25 hours, 4.3.

4.4.2 Bifurcation analysis and Results

For this second topology, bifurcation diagrams are reported in Figures 4.19, 4.20, 4.21: when I varied the parameter K_{UAS}, also the oscillatory region varies, and in particular, when K_{UAS} is higher, the oscillatory region is larger too. Moreover, when comparing this bifurcation diagrams with the diagram reported in Figure 4.7, it can be concluded that, in this case, the region of oscillatory behaviour is smaller than the same region for the system of Topology 1.

Looking at simulations and bifurcation analysis, a long oscillation period can be predicted by this model for the second topology. So, the parameters that can make this period smaller were studied by using bifurcation diagram: some diagrams of period are reported as function of the new parameters from Figure 4.22 to Figure 4.23, and in Figure 4.24 a bifurcation diagram between two significant degradation rate is reported: from these data, we can observe that the period can vary from 729 to 9260 minutes, with d_3 and d_{PG} varying respectively in the intervals $[0.0009, 0.0198]$ and $[0.0007, 0.015]$. This implies that changing the degradation of the miR223 protein can introduce a smaller period in the system.

Table 4.1: Parameters from literature for the PNFL.

Parameters	Definition	unit	Value
v_1	maximal transcription rate for CMV promoter	$[nMmin^{-1}]$	0.075432
v_2	translation rate for tTA protein	$[min^{-1}]$	0.027131449
v_3	maximal transcription rate for CMVTET promoter	$[nMmin^{-1}]$	0.075432026
v_4	translation rate for mcherry	$[min^{-1}]$	0.0271
d_1	degradation rate for tTA mRNA	$[min^{-1}]$	0.01012906
d_2	degradation rate for tTA protein	$[min^{-1}]$	0.010016646
d_3	degradation rate for miR223 mRNA	$[min^{-1}]$	0.054
d_4	degradation rate for mcherry protein	$[min^{-1}]$	0.003236
d_5	degradation rate for d2EYFP	$[min^{-1}]$	0.00048135
δ	degradation rate for miR223 mRNA	$[min^{-1}]$	0.06
α_1	basal activity for CMVTET promoter		0.00011292
λ	maximal rate of silencing	$[min^{-1}]$	0.0081
K_1	Hill constant for miR223 equation	$[nM]$	4.807645104
K_3	Hill constant for d2EYFP equation	$[nM]$	80
K_D	folding rate for miR223	$[nM]$	0.0008
h_1	Hill constant for miR223 equation		3.163
h_3	Hill constant for d2EYFP equation		4.47

Table 4.2: Parameter for the third system with undamped oscillations: refined parameters.

Parameters	Definition	unit	Value
v_1	maximal transcription rate for CMV/CMVTET promoter	$[nMmin^{-1}]$	0.075432
v_2	translation rate for tTA protein	$[min^{-1}]$	0.027131449
v_4	translation rate for mcherry protein	$[min^{-1}]$	0.0271
d_1	degradation rate for tTA mRNA	$[min^{-1}]$	0.01012906
d_2	degradation rate for tTA protein	$[min^{-1}]$	0.010016646
d_3	degradation rate for inactive miR223 mRNA	$[min^{-1}]$	0.0004814
d_4	degradation rate for mcherry protein	$[min^{-1}]$	0.003236
d_5	degradation rate for d2EYFP protein	$[min^{-1}]$	0.00048135
δ	degradation rate for active miR223 mRNA	$[min^{-1}]$	0.007
α_1	basal activity for CMVTET promoter		0.015
λ	maximal rate of silencing	$[min^{-1}]$	0.073879
K_1	Hill constant for CMVTET promoter	$[nM]$	3
K_3	Hill constant for miR223 post transcriptional inhibition	$[nM]$	2
K_D	folding rate for miR223	$[nM]$	0.025
h_0	Hill coefficient for Doxycycline inhibition		1.62
h_2	Hill coefficient for CMVTET promoter		2
h_3	Hill coefficient for miR223 post transcriptional inhibition		4

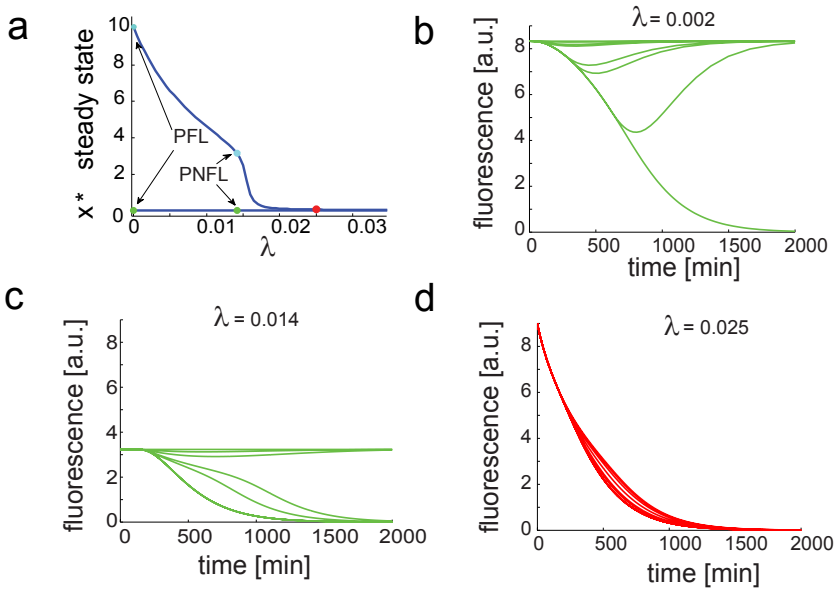


Figure 4.14: **Simulated "switch off" time-course across the PNFL cell population following different impulses of Doxycycline and with different miRNA strengths (λ).(A)** Bifurcation diagram for PNFL model with respect to parameter λ. When $\lambda = 0$ or $\lambda = 0.014$, the model represents respectively the PFL and the PNFL; in both cases the systems are bistable (green dot = OFF state, red dot = ON state). For high values of λ, the system is no more bistable and the two equilibria collapse (red point). **(B) - (C)** Different switch off time-courses were simulated by varying the parameter λ representing the strength of miRNA mediated degradation of the tTA mRNA; the value of λ can tune the dynamics of the PNFL by making it faster (stronger repression) or slower (weaker repression). (D) The same simulation is proposed for $\lambda = 0.025$ when the system is not bistable.

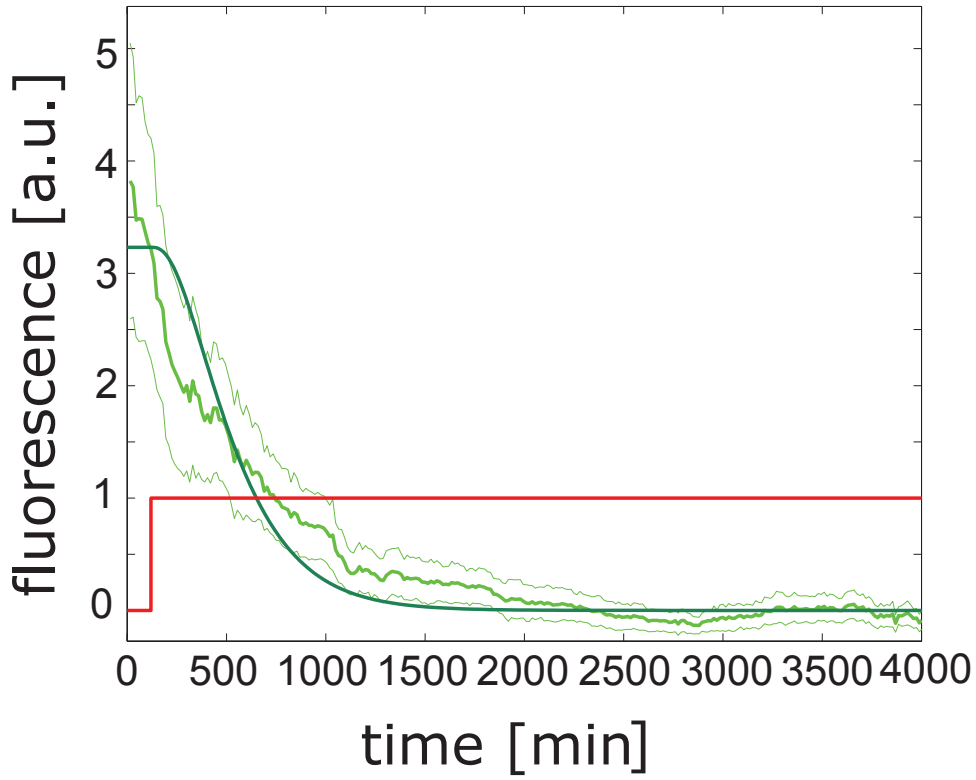


Figure 4.15: **Fitting of the PNFL parameters.** A switch off experiment was performed to estimate the unknown parameters of the PNFL. The red line represents the action of Doxycycline; the solid bright green line represents the mean fluorescence estimated on 23 replicates; the thin lines represents standard deviation, while the dark green line represents the best fit obtained with the PNFL ODE model.

Table 4.3: Parameters for the PNFL system.

Parameters	Definition	unit	Value
G_1	insertion point for the first construct		1.058
G_2	insertion point for the second construct		1
v_1	maximal transcription rate for CMVTET promoter	$[nMmin^{-1}]$	0.075432
v_2	translation rate for tTA protein	$[min^{-1}]$	0.027131449
v_3	maximal transcription rate for CMVTET promoter	$[nMmin^{-1}]$	0.075432026
v_4	translation rate for mCherry	$[min^{-1}]$	0.0271
d_1	degradation rate for tTA mRNA	$[min^{-1}]$	0.01012906
d_2	degradation rate for tTA protein	$[min^{-1}]$	0.010016646
d_3	degradation rate for miR223 mRNA	$[min^{-1}]$	0.0012351
d_4	degradation rate for d2EYFP protein	$[min^{-1}]$	0.00453
d_5	degradation rate for mCherry protein	$[min^{-1}]$	0.00048135
δ	degradation rate for miR223 mRNA	$[min^{-1}]$	0.00175
α_1	basal activity for CMVTET promoter		0.000011292
λ	maximal rate of silencing	$[min^{-1}]$	0.0142
K_1	Hill constant for miR223 equation	$[nM]$	4.807645104
K_3	Hill constant for d2EYFP equation	$[nM]$	3.97
K_D	folding rate for miR223	$[nM]$	0.0142
h_1	Hill constant for miR223 equation		3.163
h_3	Hill constant for d2EYFP equation		4

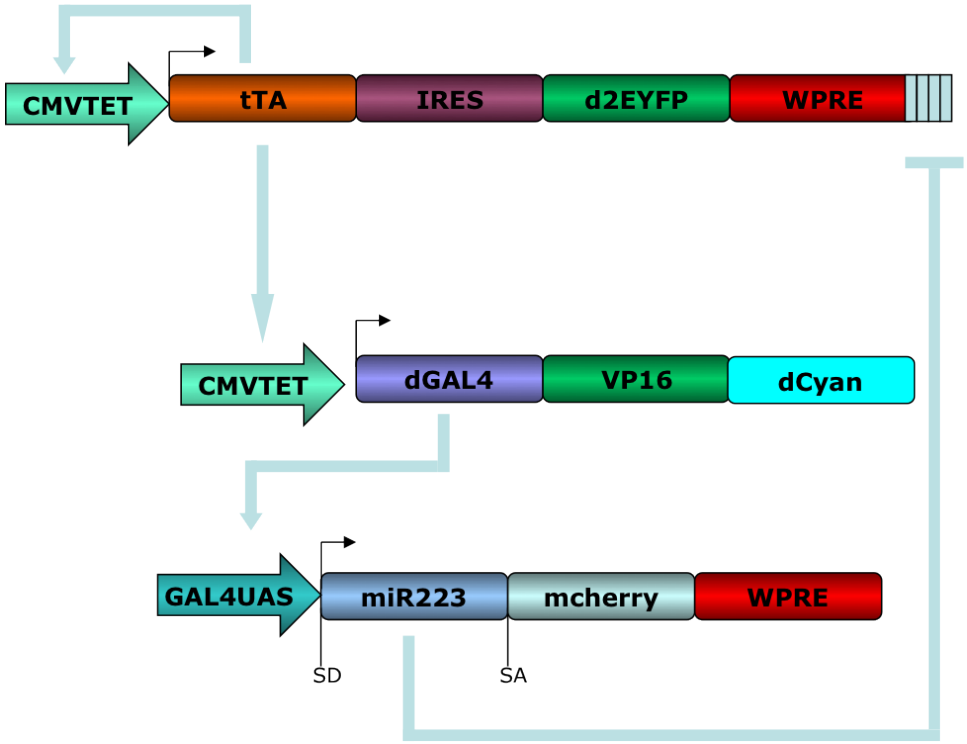


Figure 4.16: **Topology 2.** A new proposed Positive-Negative-Feedback-Loop able to present oscillations. In addition to the first topology, another step is added in the negative feedbak, thus to delay the negative interaction and amplifying the region of oscillatory behaviour.

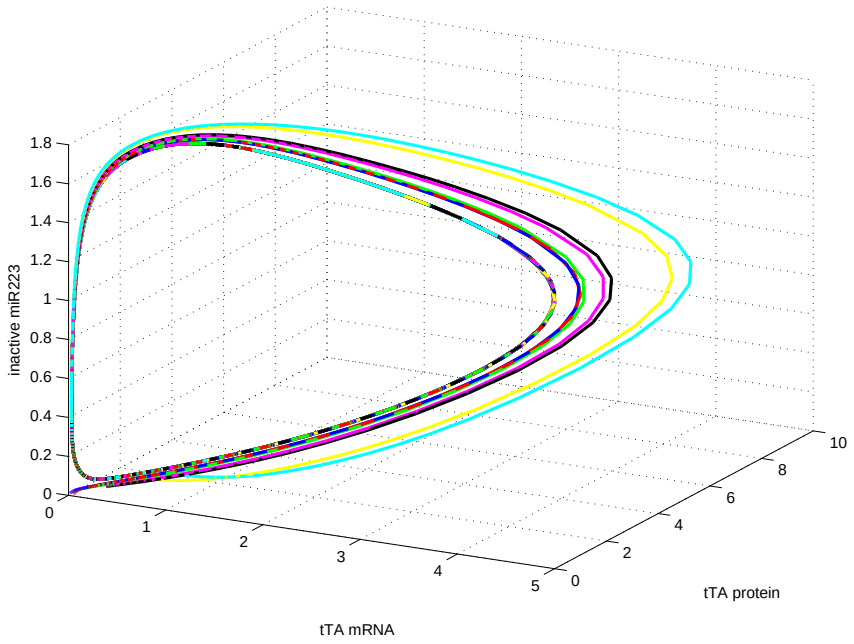


Figure 4.17: **Phase space for the model with parameters reported in Table 4.4, second topology.** A simulation in the phase space of the solutions projected on the first three components of the model reveals that a limit cycle exists for this set of parameters, thus implying the possibility of oscillations.

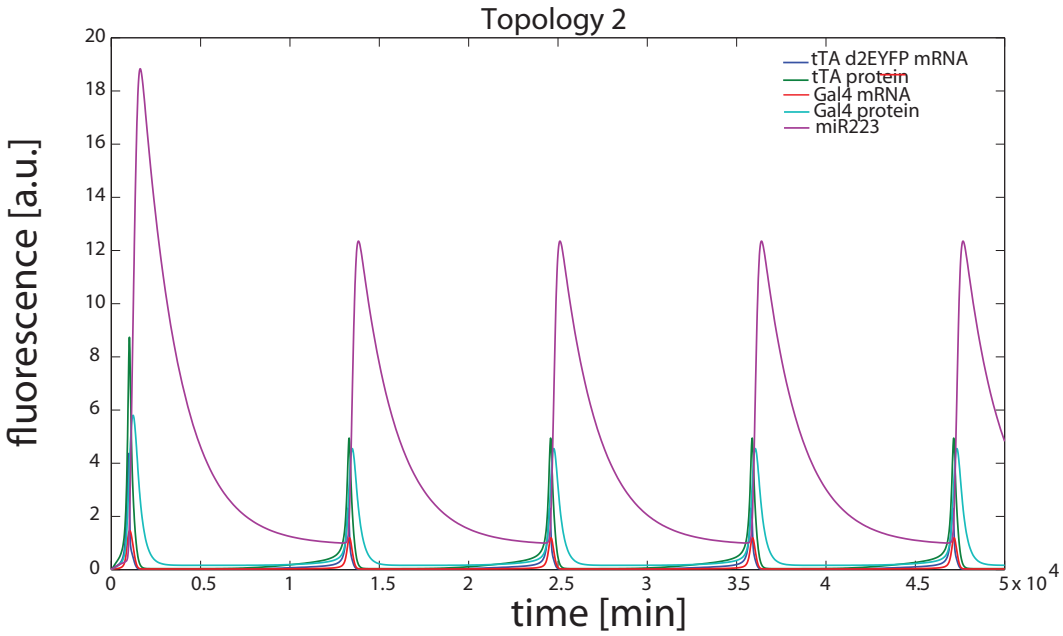


Figure 4.18: **Simulation for the oscillator of the second topology.** Simulated time course of the solutions. With respect to the first topology, the oscillations present two different time scale, going up faster than slowing down.

Table 4.4: Parameter for the first model with undamped oscillations: Topology 2.

Parameters	Definition	unit	Value
v_1	maximal transcription rate for CMV promoter	$[nMmin^{-1}]$	0.075432
v_2	translation rate for tTA protein	$[min^{-1}]$	0.027131449
v_3	maximal transcription rate for CMVTET promoter	$[nMmin^{-1}]$	0.075432026
v_4	translation rate for mcherry	$[min^{-1}]$	0.0271
d_1	degradation rate for tTA mRNA	$[min^{-1}]$	0.01012906
d_2	degradation rate for tTA protein	$[min^{-1}]$	0.010016646
d_3	degradation rate for miR223 mRNA	$[min^{-1}]$	0.0004814
d_4	degradation rate for mcherry protein	$[min^{-1}]$	0.003236
d_5	degradation rate for d2EYFP	$[min^{-1}]$	0.00048135
α_1	basal activity for CMVTET promoter		0.015
λ	maximal rate of silencing	$[min^{-1}]$	0.073879
K_1	Hill constant for miR223 equation	$[nM]$	3
K_3	Hill constant for d2EYFP equation	$[nM]$	2
h_2	Hill coefficient for miR223 equation		2
h_3	Hill coefficient for d2EYFP equation		4
α_{UAS}	Basal activity for Gal4UAS promoter		0.008
K_{UAS}	Hill constant for Gal4UAS promoter	$[nM]$	3
h_{UAS}	Hill coefficient for Gal4UAS promoter		4
d_{PG}	degradation rate for dGal4 protein	$[min^{-1}]$	0.003086
d_G	degradation rate for dGal4 mRNA	$[min^{-1}]$	0.0458
v_{UAS}	maximal transcription rate for Gal4UAS promoter	$[nMmin^{-1}]$	0.055
v_{PG}	translation rate for dGal4 protein		0.02

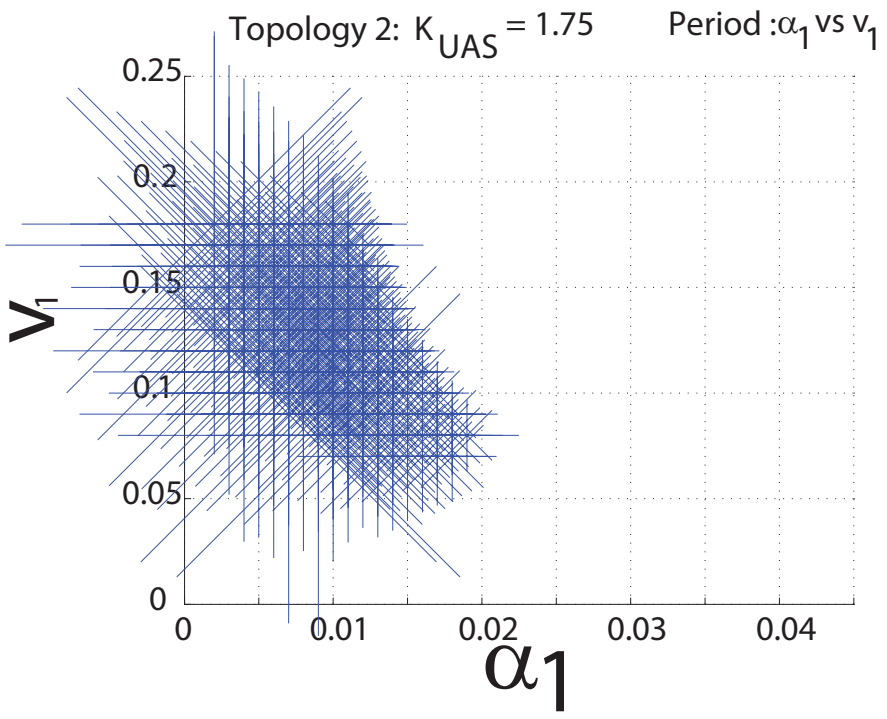


Figure 4.19: **Bifurcation diagram for the second topology between parameters α_1 and v_1, when $K_{UAS} = 1.75$.** Two variables bifurcation diagram between the leakiness and maximal transcription rate of the CMV-TET promoter, when the Michelis Menten constant of the GAL4-UAS promoter is fixed to 1.75: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

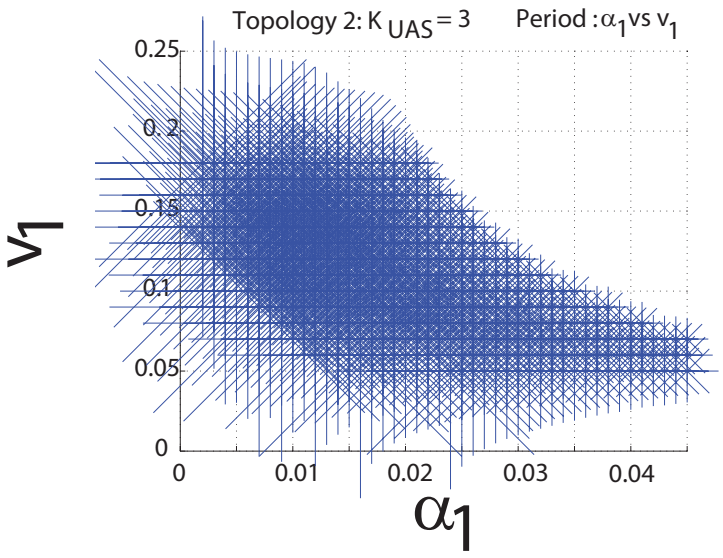


Figure 4.20: **Bifurcation diagram for the second topology between parameters α_1 and v_1, when $K_{UAS} = 3$.** Two variables bifurcation diagram between the leakiness and maximal transcription rate of the CMV-TET promoter, when the Michelis Menten constant of the GAL4-UAS promoter is fixed to 3: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

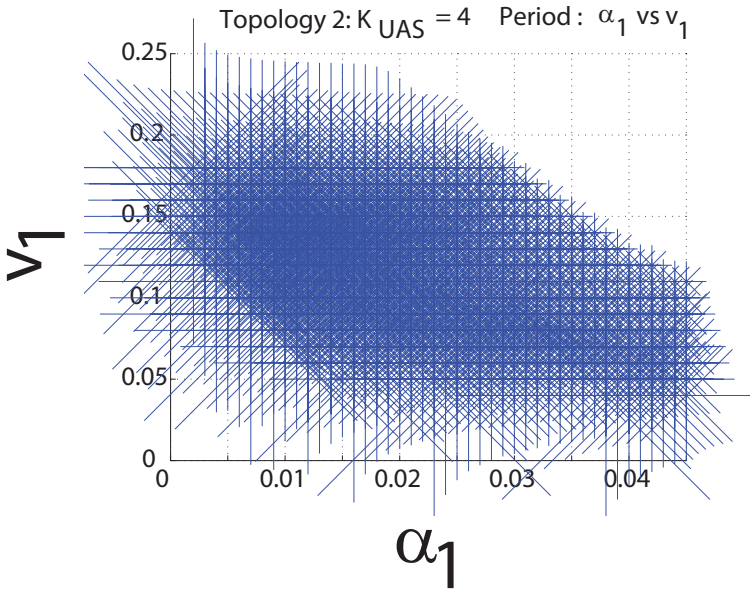


Figure 4.21: **Bifurcation diagram for the second topology between parameters α_1 and v_1, when $K_{UAS} = 4$.** Two variables bifurcation diagram between the leakiness and maximal transcription rate of the CMV-TET promoter, when the Michelis Menten constant of the GAL4-UAS promoter is fixed to 4: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

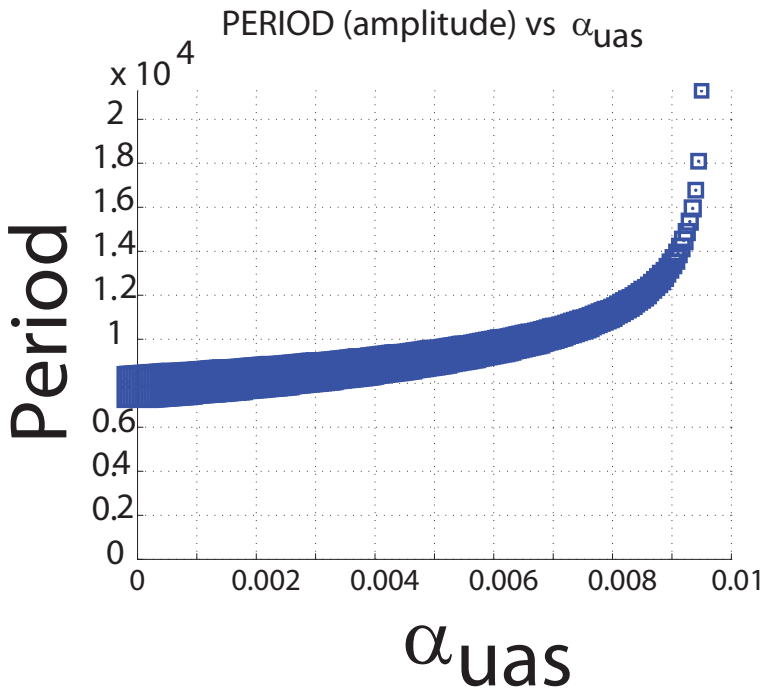


Figure 4.22: **Diagram between Period and parameter a_{UAS}.** The period increases as the paramter increases. Larger square dimensions correspond to larger amplitudes of oscillations.

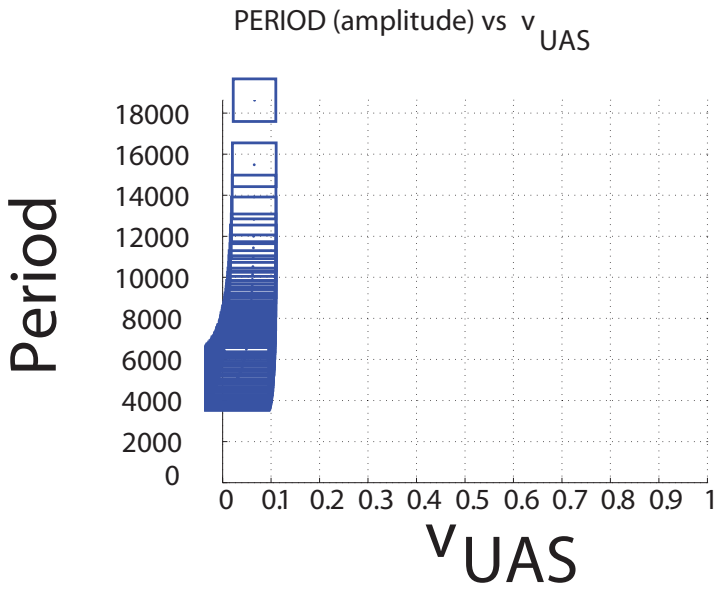


Figure 4.23: **Diagram between Period and parameter v_{UAS}.** The period increases as the paramter increases. Larger square dimensions correspond to larger amplitudes of oscillations.

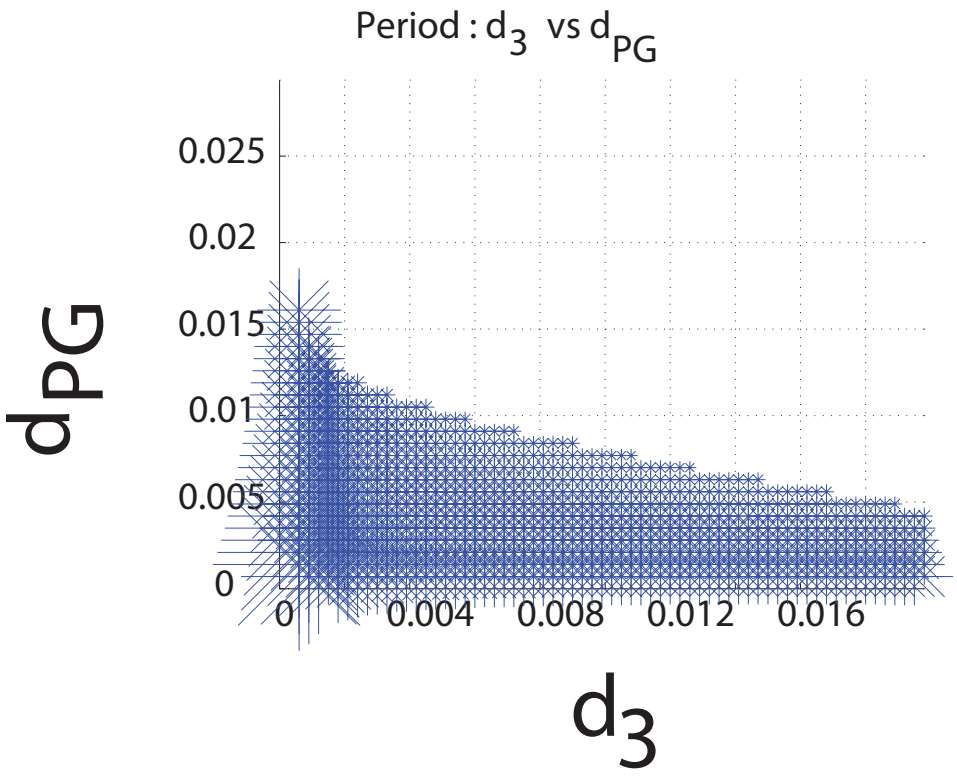


Figure 4.24: **Bifurcation diagram between parameters d_3 andd_{PG}.** Two variables bifurcation diagram between the degradation rate of the microRNA and the degradation rate of the dGal4 protein: the blue stars indicate the region in which the system presents oscillations and the greater are the stars, higher is the period.

Chapter 5

miRNA confers phenotypic robustness to the toggle-switch by buffering noise in protein expression

In this chapter, a study of the robustness of the PNFL system described in Chapter 4 is proposed.

Negative feedback loops have been widely studied in literature for their role in buffering biological noise. Here two different mathematical analyses are performed, and experimentally validated, to demonstrate this properties in-vivo in mammalian cells. Part of this chapter was published in [125].

5.1 Introduction

miRNAs are small non-coding RNAs able to modulate protein expression by pairing to cognate sites at the 3' untraslated region (UTR) of target mRNAs, thus promoting mRNA degradation and/or translation inhibition. The role of miRNAs is still under debate because, unlike transcription-factors, the phenotype induced by their perturbation (i.e. knock-down or overexpression) is usually subtle [68]. For example, in *C. elegans*, systematic knockout of single miRNAs, or even double or triple mutants, fails to elicit an observable phenotype [2, 7, 92].

It has been postulated that miRNAs play a pivotal role in conferring robustness to biological processes against enviromental fluctuations during development [44, 66, 81] and, more generally, in buffering fluctuations in gene expression [21, 29, 44, 62, 66, 68, 81, 82, 141, 147].

Biological robustness has been defined as the ability to correctly perform a biological function against endogenous and exogenous perturbations [75], such as stochastic fluctuations in gene expression. This "biological noise" has long been observed in both prokaryotes and eukaryotes, where it increases phenotypic variability among individual cells [10], [70], [46], [107]. Increased cell-to-cell variability has been found to be either benificial, such as during stress or differentiation, or detrimental such as during embryonic development of multicellular organisms, where it can decrease robustness if not properly controlled [115].

For example, correct development of sensory organs in Drosophila depends on the presence of the microRNA miR-7, but only in the face of environmental perturbations [81]. Indeed, miR-7-mutant animals subjected to fluctuating environmental temperature fail to form complete antennae, whereas wild-type larvae develop normally [81].

So far, the ability to elucidate the role of miRNAs in conferring robustness to a biological process has been hampered by the difficulty in clearly distinguishing the contribution of the miRNA from the cross-talk of all the other regulatory pathways that control the same biological process. Indeed, miRNA-mediated negative feedback loops are often found coupled to multiple transcriptional feedback loops in gene regulatory networks across species and collaborate with transcription factors to regulate their targets [81, 87, 141].

This is a complementary approach to the investigation of miRNAs in an endogenous setting and it can uncover the design principles underlying biological robustness.

5.2 Stochastic models: robustness of a Positive and microRNA-based Negative Feedback Loop system

It has been shown in Chapter 4 that the PNFL network can behave as a toggle switch. Here I will analyse its capacity to mantain ita state indefinitely

against environmental perturbation of biological noise. Ideally, once a toggle-switch is OFF (or ON), it should stay OFF (or ON) indefinitely: so an ideal toggle-switch has to maintain the ON (high protein level) or OFF (low protein level) state indefinitely across multiple cell divisions and to switch only when a specific signal is transiently applied to it. However, due to the auto-catalytic nature of both the PFL and PNFL motifs (i.e. a transcription factor self-activating its own promoter), it should be much easier for both motifs to stay ON rather than to stay OFF: this is confirmed by the fact that when the switch is OFF, small fluctuations in the transcription factor protein levels may be amplified by the positive feedback loop, causing the system to switch back ON. [70]

First, an evaluation of the phenotypic robustness of both the PNFL and PFL motifs was performed, in terms of how long they were able to stay OFF in the absence of Doxycycline.

A long-term experiment was performed by treating monoclonal population of PNFL cells and their matched control population of PFL cells with Doxycycline for 72 hrs (4320 min); after washing Doxycycline out, and d2EYFP fluorescence is monitored in cells for 96 hrs (5760 min) by Fluorescence Activated Cell Sorting (FACS). If both motifs exhibit a robust phenotype, then they should not re-express d2EYFP and stay permanently OFF.

As shown in Figure 5.1a for the PFL 7 clone and the matched PNFL

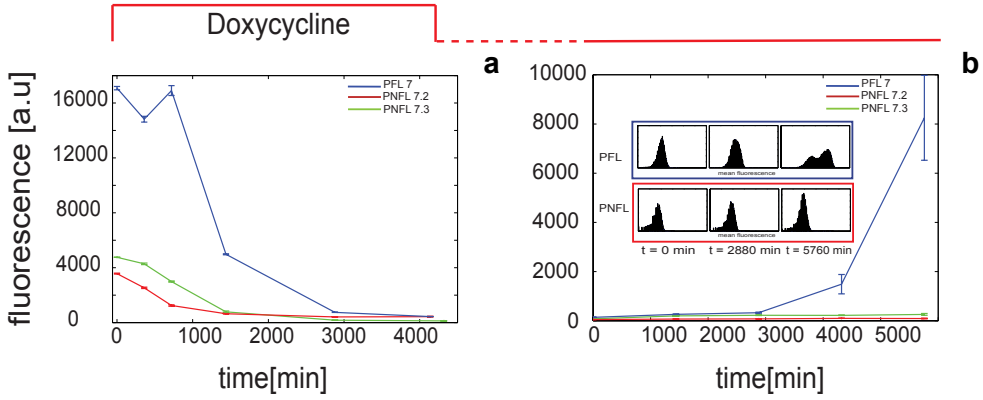


Figure 5.1: **FACS switch off experiment. (A)**: d2EYFP fluorescence levels in PFL cells (blue line) and PNFL 7-2 and 7-3 cells (respectively red - green line) were measured at 0 hrs, 6 hrs (360 min), 12 hrs (720 min), 24 hrs (1440 min), 48 hrs (2880 min) and 72 hrs (4320 min) following treatment with Doxycycline (1 μg/ml) at time 0 hrs. **(B)**: d2EYFP fluorescence levels in PFL cells (blue line) and PNFL cells (red line) cells were measured at 0 hrs, 24 hrs (1440 min), 48 hrs (2880 min), 96 hrs (5760 min) following removal of Doxycycline at time 0 hrs. Prior to time 0 hrs, both PFL and PNFL cells were grown in the presence of Doxycycline for 72 hrs. (Subpanel): histogram displaying FACS data for PFL 7 (blue frame) and PNFL 7-2 (red frame) at times $t = 0$, 2880 min and 5760 min; CVs computed at 0 hrs, 24 hrs, 48 hrs, 72 hrs are respectively 74.5, 105.1, 108, 202.6 for PFL and 81.9, 88.4, 85.75, 96 for the PNFL, thus revealing the higher variability for PFL clones compared to PFL.

7-2 and 7-3 clones, following Doxycycline treatment, both the PNFL and PFL cells switched d2EYFP expression OFF, with dynamics similar to those estimated using the microfluidics setup.

Strikingly, following Doxycycline washout, PFL cells started switching back ON at 72 hrs (4320 min) and fully expressed d2EYFP at 96 hrs (5670 min) (Figure 5.1b). On the contrary, PNFL cells stayed OFF (Figure 5.1b).

As further confirmation of the miRNA role, the same experiment was performed on PNFL cells but this time inhibiting the action of the miRNA by transfection with a Locked Nucleic Acid complementary to the miR223 sequence (LNA 223) [109]. PNFL cells treated with the LNA increased d2YFP expression at 72 and 96 hrs, whereas PNFL cells treated with a control LNA maintained their OFF state, confirming our previous results (Fig. 5.6).

This long-term experiment was repeated on additional PFL and matched PNFL clones, in order to exclude that the difference in phenotypic robustness I observed was clone-specific. As reported in Figure 5.2a,b for the PFL 2 and PNFL 2-2, 2-3 and 2-7 clones, and in Figure 5.2c,d for the PFL 10 and PNFL 10-2 clones, the presence of the miRNA-mediated feedback loop consistently confers robustness to the PFL motif, allowing the PNFL cells to maintain their OFF state for longer as compared to the matched PFL cells. The main difference among the three PFL clones (PFL-7, PFL-2 and PFL-10) was observed in the time taken to switch back ON following Doxycyline removal (Fig. 5.1b and Figure 5.2b,d). This difference is likely due to the different

insertion site of the PFL cassette in the genome of the three clones.

In order to analyse theoretically the effect of the microRNA on noise propagation in the PNFL network motif, a simplified stochastic gene expression model has been developed with transcription, translation and degradation as probabilistic events occurring at exponentially distributed time intervals [104, 128].

Specifically, in the model, mRNAs are produced in transcriptional bursts from the promoter, in agreement with the recent modeling and experimental evidences concerning noise in eukaryotic gene expression [10, 20, 70].

Under a set of simplifying assumptions, a system of differential equations describing the dynamics of the statistical moments in four different scenarios (Fig. 5.7) can be derived: (1) unregulated transcription, i.e. the mRNA is transcribed from a constitutive promoter [128]; (2) regulated transcription where the protein activates its own transcription (i.e. the PFL motif); (3) a negative feedback loop mediated by the microRNA inducing degradation of the target mRNA (i.e. a NFL motif); (4) the PNFL motif.

Finally, an analytical expression for the Coefficient of Variation (CV) was computed, i.e. the relative deviation of protein expression in each cell compared to the population average, which can be used as a measure of noise. Indeed, a small CV corresponds to a tight distribution centered around the mean, hence a small cell-to-cell variability; a large CV corresponds to a loose distribution, indicating large cell-to-cell variability.

5.2.1 Simplified stochastic models of NOPFL, PFL and PNFL motifs

In order to investigate the effect of the miRNA on the protein noise level, the model of gene expression as shown in Figure 5.7 was used: in this model, the mRNA is assumed to be transcribed in bursts and proteins are translated from single mRNA molecules [128]. In this simplified model it is possible to derive differential equations describing the time evolution of the different statistical moments of the mRNA and protein counts [128]. From these moments, one can derive an analytical expression for the steady-state protein noise level CV^2 = (Squared Coefficient of Variation) as function of the mean number of mRNA molecules ($\overline{m}$) for the four configurations in Figure 5.7 following the method derived by [128].

Following [128], the Squared Coefficient of Variation CV^2 for a simple transcription- translation model (NOPFL in Fig. 5.7A) can be written as:

$$CV^2{}_{NOPFL} = \frac{\gamma_p(< B > + < B^2 >)}{2 < B > (\gamma_p + \gamma_m) < \overline{m} >}, \tag{5.1}$$

where γ_m and γ_p represent the degradation rates for the mRNA and the corresponding protein, B is the number of mRNA molecules produced per burst of transcription, $\overline{m}$ is the mean number of mRNA molecules at the steady state.

In the case of regulated transcription, where the protein activates its own transcription, as depicted in Fig. 5.7B (PFL), the following expression can

be derived:

$$CV^2{}_{PFL} = \frac{\gamma_p(< B > + < B^2 >)}{2 < B > (\gamma_p + \gamma_m)(1 - \overline{k}) < \overline{m} >}; \tag{5.2}$$

where $\overline{k}$ (a positive number) represents the effect of the protein on the mRNA transcription rate. By comparing Eq.5.1 to Eq.5.2 it is immediately clear that $CV^2{}_{PFL}$ is always greater than $CV^2{}_{NOPFL}$, that is the PFL increases noise levels.

In order to derive a simplified expression for CV^2 in the presence of a miRNA, I assumed that the miRNA is transcribed together with the mRNA as represented in Figure 5.7C and that the degradation rate of the mRNA is proportional to the number of miRNA molecules, which in turn are proportional to the number of mRNA molecules, thus yielding the following equation for the mRNA degradation rate:

$$\gamma_m(m) = \left(1 - k\frac{m - < \overline{m} >}{< \overline{m} >}\right)\overline{\gamma_m}. \tag{5.3}$$

We can now write the following set of differential equations for the Negative Feedback Loop in Figure 5.7C:

$$\frac{d<m>}{dt} = k_m < B > - < \gamma_m m > \tag{5.4}$$

$$\frac{d<p>}{dt} = k_p < m > - \gamma_p < m > \tag{5.5}$$

$$\frac{d<m^2>}{dt} = k_m < B^2 > + 2k_m < B >< m > + < \gamma_m(m)m > - 2 < \gamma_m(m)m^2 > \tag{5.6}$$

$$\frac{d<mp>}{dt} = k_m < B >< p > - < \gamma_m mp > + k_p < m^2 > - \gamma_p < mp > \tag{5.7}$$

$$\frac{d<p^2>}{dt} = k_p < m > + 2k_p < mp > + \gamma_p < p > - 2\gamma_p < p^2 > \tag{5.8}$$

$$\tag{5.9}$$

As before, denoting the number of mRNA and protein molecules at steady state by $< \overline{m} >$ and $< \overline{p} >$, I obtain:

$$CV^2{}_{NFL} = \frac{\gamma_p(< B > + < B^2 >)}{2 < B > (\gamma_p + \gamma_m(1 + \overline{\overline{k}})) < \overline{m} >}. \tag{5.10}$$

where the parameter $\overline{\overline{k}}$ is a positive real number representing the effect of the miRNA on mRNA degradation.

By comparing Eq.5.2, Eq.5.1 and Eq.5.10, it is possible to check that that the following relationship is always satisfied: $CV^2{}_{NOPFL} \leq CV^2{}_{NFL} \leq CV^2{}_{PFL}$. Hence the effect of miRNA acting on mRNA degradation is to decrease noise in protein levels.

The case of a transcriptional PFL coupled to a miRNA-mediated NFL,

as shown in Figure 5.7, is less straightforward and it is harder to derive an analytical expression for CV^2, hence I formulated an heuristic approximation for the PNFL by analogy to the previous results:

$$CV^2{}_{PNFL} = \frac{\gamma_p^*(< B > + < B^2 >)}{2 < B > (1 - \overline{k})(\gamma_p^* + \gamma_m^*(1 + \overline{\overline{k}})) < \overline{m} >} \tag{5.11}$$

The value of $CV^2{}_{PNFL}$ now depends both on the strength of the positive feedback loop mediated by the protein $\overline{k}$ and on the strength of the negative feedback loop $\overline{\overline{k}}$ mediated by the miRNA.

Figure 5.3 shows a graph of the protein noise level (in red) for the PNFL motif ($CV^2{}_{PNFL}$) as a function of the mRNA expression level ($\overline{m}$) compared to the protein noise level (in blue) of the PFL motif ($CV^2{}_{PFL}$) rescaled and fitted to the experimental data obtained by FACS analysis of the PFL and PNFL clones.

The numerical values of the parameters in Eq.5.2 and Eq.5.11 used in Figure 5.3 were derived as follows: γ_m and γ_p correspond to the parameters d_1 and d_2 in Table 4.3, whereas $< B >$ and $< B^2 >$ were chosen to minimize the Mean Squared Error to the experimental FACS data in Figure 5.3. Table 5.1 lists the values of the fitted parameters.

From the analytical expression of the CV in the four scenarios, it can be deduced that, independently of the model parameter values, noise increseas when a transcriptional positive feedback loop is present, compared to unregulated transcription, in agreement with previous modelling and experimental

results [10, 17, 70]. On the contrary, noise decreases, compared to unregulated transcripiton, when a microRNA-mediated negative feedback loop is present, again in agreement with previous modelling and experimental results on negative feedbacks [10, 70, 118]. In the case of the PNFL motif, when both negative and positive feedbacks are present, noise is reduced compared to the PFL motif, but is greater than the unregulated transcription.

Table 5.1: Parameters for the stochastic model of CV^2 as represented in Figure 7 of the main text.

Parameters	numeric value
$< B >$	0.5595
$< B^2 >$	6.1434
k	0.7508
$\overline{k}$	0.6077

5.2.2 Stochastic models of the PFL and PNFL noisy behaviour: Master Equations

The advantage of the simplified model, presented in the previous section, is that it allows to compare analytically the CV^2 for the different motifs; the disadvantage is that the model maybe too simplified thus yielding unrealistic results.

We therefore also derived two master equations, one for the PNFL and

one for the PFL motif, which take into account the discrete and stochastic nature of chemical reactions [104, 137].

In order to write the PNFL master equation, I generalized the deterministic model described by Eqs.4.1 and disregarded equations describing the dynamics of the reporter proteins.

By denoting the tTA mRNA, tTA protein and miR223 mRNA with the variables $w \triangleq x_1$, $q \triangleq x_2$ and $r \triangleq x_4$ and assuming quasi-steady state for Eq.(6) so that $x_3 \propto x_4$, the deterministic model for the PNFL (Eqs.4.1) becomes:

$$\frac{dw}{dt} = (\overline{K_w} + K_w(q)) - g_w w - \beta_w(r) w \tag{5.12}$$

$$\frac{dq}{dt} = K_q w - g_q q \tag{5.13}$$

$$\frac{dr}{dt} = (\overline{K_r} + K_r(q)) - g_r r \tag{5.14}$$

where $\overline{K_w} = G_1 v_1 \alpha_1$, $K_w(q) = G_1 v_1 (1 - \alpha_1) \frac{q^{h_2}}{K_1^{h_2} + q^{h_2}}$, $g_w = d_1$, $\beta_w(r) = \lambda \frac{r^{h_3}}{{K_3}^{h_3} + r^{h_3}}$, $K_q = v_2$, $g_q = d_2$, $\overline{K_r} = G_2 v_1 \alpha_1$ and $K_r(q) = K_w(q)$.

Similarly, the PFL model in [126] can be rewritten as:

$$\frac{dw}{dt} = (\overline{K_w} + K_w(q)) - g_w w \tag{5.15}$$

$$\frac{dq}{dt} = K_q w - g_q q \tag{5.16}$$

Following the method described in [104], I linearized the nonlinear functions $K_w(q)$, $K_r(q)$ and $\beta_w(r)$ to obtain:

$$K_w(q) \simeq {K_w}^0 + {K_w}^1 q \tag{5.17}$$

$$K_r(q) \simeq {K_r}^0 + {K_r}^1 q \tag{5.18}$$

$$\beta_w(r) \simeq {\beta_w}^0 + {\beta_w}^1 r \tag{5.19}$$

where ${K_w}^0$,${K_w}^1$,${K_r}^0$,${K_r}^1$,${\beta_w}^0$ and ${\beta_w}^1$ are constant terms.

The Master equation for the PFL circuit can be derived:

$$\begin{aligned}\frac{\partial P_{w,q}}{\partial t} &= (\overline{K_w} + K_w(q))(P_{w-1,q} - Pw, q) + \\ &K_q w(P_{w,q-1} - Pw, q) + g_w((w+1)P_{w+1,q} - wP_{w,q}) \\ &+ g_q((q+1)P_{w,q+1} - qP_{w,q})\end{aligned} \tag{5.20}$$

Introducing the moment generating function,([104], [137]):

$$F(z_1, z_2) = \sum_{w,q} {z_1}^w {z_2}^q P_{w,q}, \tag{5.21}$$

it is possible to derive the following Partial Differential Equation (PDE) equation:

$$\begin{aligned}\partial_t F &= (\overline{K_w} + {K_w}^0)F(z_1 - 1) + \\ &+ {K_w}^1 z_2 \partial_{z_2} F(z_1 - 1) + \\ &+ K_q z_1 \partial_{z_1} F(z_2 - 1) + \\ &+ g_w \partial_{z_1} F(1 - z_1) + g_q \partial_{z_2} F(1 - z_2)\end{aligned} \tag{5.22}$$

which at the steady state becomes:

$$(\overline{K_w} + {K_w}^0)F(z_1 - 1) + {K_w}^1 z_2 \partial_{z_2} F(z_1 - 1) +$$
$$+ K_q z_1 \partial_{z_1} F(z_2 - 1) + g_w \partial_{z_1} F(1 - z_1) + g_q \partial_{z_2} F(1 - z_2) = 0 \tag{5.23}$$

From this equation, I can estimate the moments of the distribution in order to evaluate the $CV = \frac{\sigma_q}{<q>} = \frac{\sqrt{F_{22} - {F_2}^2 + F_2}}{F_2}$.

The moments are listed below:

$$F_1 = \frac{g_q(\overline{K_w} + {K_w}^0)}{(g_w g_q - {K_w}^1 K_q)} \tag{5.24}$$
$$F_2 = \frac{K_q F_1}{g_q} \tag{5.25}$$
$$F_{12} = \frac{g_q F_2[(\overline{K_w} + {K_w}^0 + {K_w}^1)g_w + g_q g_w + (\overline{K_w} + {K_w}^0)g_q]}{(g_w + g_q)(g_w g_q - {K_w}^1 K_q)} \tag{5.26}$$
$$F_{11} = \frac{(\overline{K_w} + {K_w}^0)F_1 + {K_w}^1 F_{21}}{g_w} \tag{5.27}$$
$$F_{22} = \frac{K_q F_{12}}{g_q} \tag{5.28}$$

where it is possible to demonstrate that F_{12} equals F_{21}.

Similarly, I derived **the Master Equation for the PNFL circuit**:

$$\frac{\partial P_{w,q,r}}{\partial t} = (\overline{K_w} + K_w(q))(P_{w-1,q,r} - P_{w,q,r}) +$$
$$+ K_q w(P_{w,q-1,r} - P_{w,q,r}) +$$
$$+ (\overline{K_w} + K_w(q))(P_{w,q,r-1} - P_{w,q,r}) +$$
$$+ g_w((w+1)P_{w+1,q,r} - wP_{w,q,r}) +$$
$$+ g_q((q+1)P_{w,q+1,r} - qP_{w,q,r}) +$$
$$+ g_r((r+1)P_{w,q,r+1} - rP_{w,q,r}) +$$
$$+ \beta_w(r)((w+1)P_{w+1,q,r} - wP_{w,q,r}) \tag{5.29}$$

Using the moment generating function:

$$F(z_1, z_2, z_3) = \sum_{w,q,r} {z_1}^w {z_2}^q {z_3}^r P_{w,q,r} \tag{5.30}$$

it is possible to derive a PDE equation:

$$\partial F_t = (\overline{K_w} + {K_w}^0)(z_1 - 1)F +$$
$$+ {K_w}^1 z_2 F_1(z_1 - 1) + K_q z_1 F_1(z_2 - 1) +$$
$$+ (\overline{K_w} + {K_w}^0)F(z_3 - 1) + {K_w}^1 z_2 F_2(z_3 - 1) +$$
$$+ g_w F_1(1 - z_1) + g_q F_2(1 - z_2) + g_r F_3(1 - z_3) +$$
$$+ {\beta_w}^0 F_1(1 - z_1) + {\beta_w}^1 z_3 F_{13}(1 - z_1) \tag{5.31}$$

which at the steady state becomes:

$$(\overline{K_w} + {K_w}^0)(z_1 - 1)F + {K_w}^1 z_2 F_1(z_1 - 1) + K_q z_1 F_1(z_2 - 1) +$$
$$+ (\overline{K_w} + {K_w}^0)F(z_3 - 1) + {K_w}^1 z_2 F_2(z_3 - 1) + g_w F_1(1 - z_1) +$$
$$+ g_q F_2(1 - z_2) + g_r F_3(1 - z_3) + {\beta_w}^0 F_1(1 - z_1) +$$
$$+ {\beta_w}^1 z_3 F_{13}(1 - z_1) = 0 \tag{5.32}$$

From this equation, it is possible to estimate the moments in order to evaluate the $CV = \frac{\sigma_q}{<q>} = \frac{\sqrt{F_{22} - {F_2}^2 + F_2}}{F_2}$:

$$F_1 = \frac{(\overline{K_w} + {K_w}^0 - {\beta_w}^0) + {K_w}^1 F_2 - {\beta_w}^1 F_3}{g_w} \tag{5.33}$$

$$F_2 = \frac{K_q F_1}{g_q} \tag{5.34}$$

$$F_3 = \frac{(\overline{K_r} + {K_r}^0) + {K_r}^1 F_2}{g_r} \tag{5.35}$$

$$F_{11} = \frac{(\overline{K_w} + {K_w}^0 - {\beta_w}^0) F_1 + {\beta_w}^1 (F_3 - F_{31}) + {K_w}^1 F_{21} + {\beta_w}^0}{g_w} \tag{5.36}$$

$$F_{12} = \frac{(\overline{K_w} + {K_w}^0 + {K_w}^1 - {\beta_w}^0) F_2 + {K_w}^1 F_{22} + K_q F_{11}}{g_w + g_q} + \\ + \frac{K_q F_1 - {\beta_w}^1 F_{32}}{g_w + g_q} \tag{5.37}$$

$$F_{13} = \frac{(\overline{K_w} + {K_w}^0 - {\beta_w}^0 - {\beta_w}^1) F_3 + (\overline{K_r} + {K_r}^0) F_1}{g_w + g_r} + \\ + \frac{{K_r}^1 F_{21} + {K_w}^1 F_{23} - {\beta_w}^1 F_{33}}{g_w + g_r} \tag{5.38}$$

$$F_{22} = \frac{K_q F_{12}}{g_q} \tag{5.39}$$

$$F_{23} = \frac{(\overline{K_r} + {K_r}^0 - {K_r}^1) F_2 + {K_r}^1 F_{22} + K_q F_{13}}{g_r + g_q} \tag{5.40}$$

$$F_{33} = \frac{(\overline{K_r} + {K_r}^0) F_3 + {K_r}^1 F_{23}}{g_r} \tag{5.41}$$

We numerically solved Eqs. (5.24 - 5.28) for the PFL and Eqs.(5.33 - 5.41) for the PFL to simulate the CV^2 as function of the mean protein level. Results are shown in Figure 5.8 and parameters used for the simulations are reported in Table 5.2. This analysis further confirms our thesis of an higher robustness of the PNFL circuit compared to the PFL in terms of protein noise levels. In the same Figure 5.8 the simulation for the NOPFL is also

reported using the moments already computed in [104] for the open circuit (TF - target) in which the parameters in common to our PFL were chosen with the same values.

Table 5.2: Parameters for the simulations reported in Figure 5.8. Note that the common parameters are described by the same values.

Parameters	numeric value
$\overline{K_w}$	1.1882
${K_w}^0$	0.143
${K_w}^1$	0.2331
$\overline{K_r}$	$=\overline{K_w}$
${K_r}^0$	$= {K_w}^0$
${K_r}^1$	$= {K_w}^1$
K_q	1.3621
g_w	0.4847
g_q	0.6557
g_r	0.3
${\beta_w}^0$	0.01
${\beta_w}^1$	0.005

5.3 Experimental validation

In order to investigate the validity of the stochastic model predicitions, a measure of the distribution of fluorescence levels across the cells in each of the monoclonal cell populations (9 for the PFL and 14 for the PNFL) has performed, as shown in Fig. 5.3a. Then the protein noise level was estimated by computing the (CV) as the relative deviation of d2EYFP fluorescence in each cell compared to the clonal population average. Figure 5.3b reports the theoretical CVs derived for both the PFL and the PNFL stochastic models, superimposed to the experimentally measured CVs in the 23 clones (9 PFL and 14 PNFL).

Figure 5.3c shows representative examples of the d2EYFP fluorescence distribution of two PFL clones and two PNFL clones, with similar average fluorescence intensities; Figure 5.3b shows the CV for all the PFL and PNFL clones.

The experimental results clearly demonstrate a reduced protein noise level (CV) in the PNFL cells as compared to the PFL cells, indicating a specific effect of the miRNA-mediated negative feedback in buffering protein expression fluctuations.

The PFL motif is predicted to have a higher noise level than the PNFL motif because fluctuations in the number of transcribed mRNAs are translated into fluctuations in the number of proteins, which in turn control the number of transcribed mRNAs. This results in a net amplification of noise.

On the other hand, when the miRNA-mediated negative feedback is present, noise is reduced because fluctuations in the number of mRNA molecules amplified by the PFL are counterbalanced by the presence of the miRNA.

5.4 Discussion

The combination of computational analysis coupled with experimental biology, at the core of synthetic biology, has been used as a powerful tool to elucidate complex cellular processes, such as cell polarization [30] and transcriptional regulation [21].

Here, I used the synthetic biology paradigm to explore the role of miRNAs in conferring robustness to a fundamental network motif, the Positive Feedback Loop.

A Positive Feedback Loop enables a gene regulatory network to encode a "cellular memory" due to its switch-like behaviour created by two different stable states [3, 69]. Transitions across the states (i.e. switching) can be triggered by a specific endogenous or exogenous event [62, 147], thus enabling the control of key decision processes such as cell differentiation and cell fate in response to specific clues [37].

The switch-like behaviour of the PFL motif, however, comes at a cost: first, the PFL may considerably slow down the transcriptional response of the circuit to the triggering event [126]; this can be acceptable, or necessary, in some cases, but detrimental in others; the second and biggest issue is that the

PFL motif increases biological noise because small fluctuations in the expression levels of the transcription factor are amplified by the self-regulatory loop. The increased cell-to-cell variability in the transcription factor concentration may cause the transcription factor level to be higher than the promoter activation threshold, inducing transition of the switch from the OFF state to the ON state, even in the absence of the triggering event [3, 62].

Here, I experimentally demonstrated that the presence of a miRNA suffices to solve both issues at once, by conferring phenotypic robustness to the toggle-switch, thanks to the reduction of biological noise, and by enabling the circuit to respond faster.

Therefore, the difference in robustness between the PNFL and PFL motifs can be explained by taking into account the "stochastic switching" phenomenon. Stochastic fluctuations in gene expression, due to biological noise, will induce random transitions between the two states of the toggle-switch [3, 90, 144].

Interestingly, if the miRNA repression on the target mRNA is too strong, then the PNFL will no longer be able to function as a toggle-switch because of a loss of bistability (4.14). This intuitively means that the miRNA will force the circuit to always be OFF by inhibiting the transcription factor expression. In addition, the stronger the miRNA repression, the smaller the difference between the OFF state and the ON state in terms of protein expression levels. These constraints may have contributed to the evolution of the weak

transcriptional effect exerted by the majority of the known miRNAs.

A recent experimental and computational analysis investigated the properties of the incoherent feedforward loop motif (FFL) consisting a transcription factor inducing expression of both a reporter gene and a miRNA directed against the same reporter gene [21]. The authors demonstrated that the FFL made the reporter gene expression level insensitive to the motif-carrying plasmid copy number. Moreover, they observed that the presence of the miRNA led to a reduced variability in reporter protein expression among cells transfected with the motif-carrying plasmid. Deans et al. used a small short-hairpin RNA (shRNA) to enhance the robustness and performance of an inducible mammalian genetic switch. [39]

Our results add further support to the role of miRNAs as noise suppressor. In addition, I demonstrated that reducing noise is essential to maintain a robust phenotype by preventing cells from unwanted stochastic switching.

We believe that our findings on the role of the miRNA explain the abundance of miRNA-mediated negative feedbacks in gene regulatory networks across species [87,141]. Indeed, it was previously suggested [141] that miRNA-mediated NFLs may be needed to reduce intracellular noise, thus conferring phenotypic robustness to the pathways they are part of. Here, I demonstrated that indeed this is the case.

One of the best studied examples of PNFL in vivo is the E2F1/miR-17-92 network [102, 135, 141]. Here the E2F1 transcription factor, a main regulator

of cell cycle, regulates its own transcription and that of the miR-17-92 cluster targeting E2F1 itself. According to our results, and as previously suggested, the role of the miRNA-mediated NFL should be of preventing stochastic activation of E2F1 leading to unwanted proliferation or apoptosis [135, 141].

In addition, several developmental transcription factors are regulated by positive feedback loops [37]. These PFLs could be frequently coupled to miRNA-mediated NFLs to prevent untimely activation of the transcription factor, which would likely be lethal for the developing embryo [141]. For example the muscle regulatory factor MYOD activates its own transcription [47], [138] as well as transcription of miRNAs predicted to target MyoD itself [135], [50].

Another example is provided by intestinal cell specification in the nematode C. elegans during embryonic development. This process is regulated by a positive feedback loop involving *elt-2*, the master regulatory gene of intestinal differentiation [115]. Raj et al [115] demonstrated that stochastic activation of the *elt-2* PFL underlies incomplete penetrance of mutant embryos, leading to pronounced phenotpyic variability.

A limitation of our synthetic network approach is that the miRNA seed sequence is perfectly complementary to the four recognition sites I engineered at the 3' *ttA_IRES_d2EYFP* of the mRNA (Fig. 4.13a). A perfectly complementary binding site causes the miRNA to induce a catalytic-like RNA interference process, rather than the more subtle gene expression "thresh-

olding" effect observed in endogenous miRNA binding sites [98]. This effect causes the miRNA-induced target mRNA repression to decrease, and eventually be completely abolished, if the target mRNA increases above a certain threshold saturating the miRNA pool [98]. Our results are relevant to endogenous miRNA regulation as long as the mRNA expression remains below the threshold since, in this situation, endogeneous miRNA binding sites and perfectly complementary binding sites have very similar effects. [98]

Our findings are a direct proof of the essential role of microRNAs in providing phenotypic robustness to a biological process. Our results are also relevant to the growing field of synthetic biology, in that they demonstrate how to synthetise robust circuits by employing miRNA-mediated feedback loops.

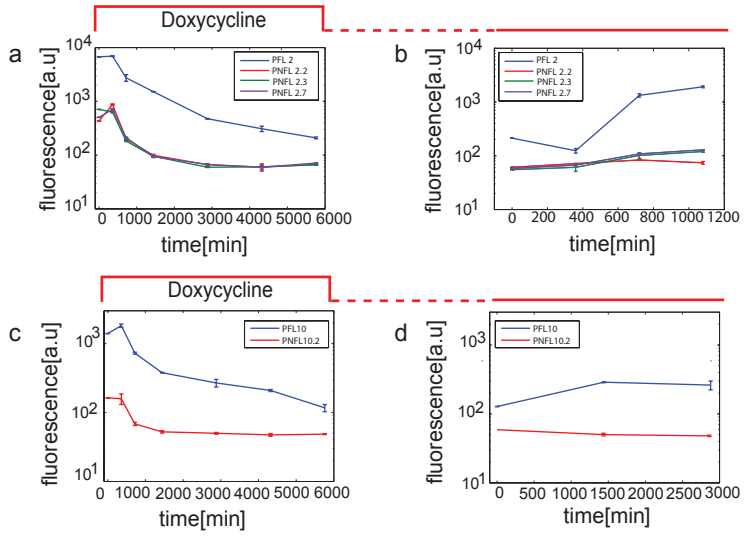


Figure 5.2: **FACS switch off experiment of PFL 2, PFL 10 cells and the corresponding PNFL cells.** PFL represented by blue line, PNFL green, red and purple.**(a)**: d2EYFP fluorescence levels in PFL 2 cells and PNFL 2-2, 2-3, 2-7 were measured at 0, 6, 12, 24, 48, 72 hrs (0, 360, 720, 1440, 2880, 4320 min) following treatment with Doxycycline (1 μg/ml) at time 0 hrs. **(b)**: d2EYFP fluorescence levels in PFL cells and PNFL cells cells were measured at 0, 6, 12, 18 hrs following removal of Doxycycline at time 0 hrs. Prior to time 0 hrs, both PFL and PNFL cells were grown in the presence of Doxycycline for 72 hrs. **(c)**: d2EYFP fluorescence levels in PFL 10 cells and PNFL 10-2 were measured at 0, 6, 12, 24, 48, 72 hrs (0, 360, 720, 1440, 2880, 4320 min) following treatment with Doxycycline (1 μg/ml) at time 0 hrs. **(d)**: d2EYFP fluorescence levels in PFL cells and PNFL cells cells were measured at 0, 24, 48 hrs (0, 1440, 2880 min) following removal of Doxycycline at time 0 hrs. Prior to time 0 hrs, both PFL and PNFL cells were grown in the presence of Doxycycline for 72 hrs. Error bars represent the standard deviation among three replicates.

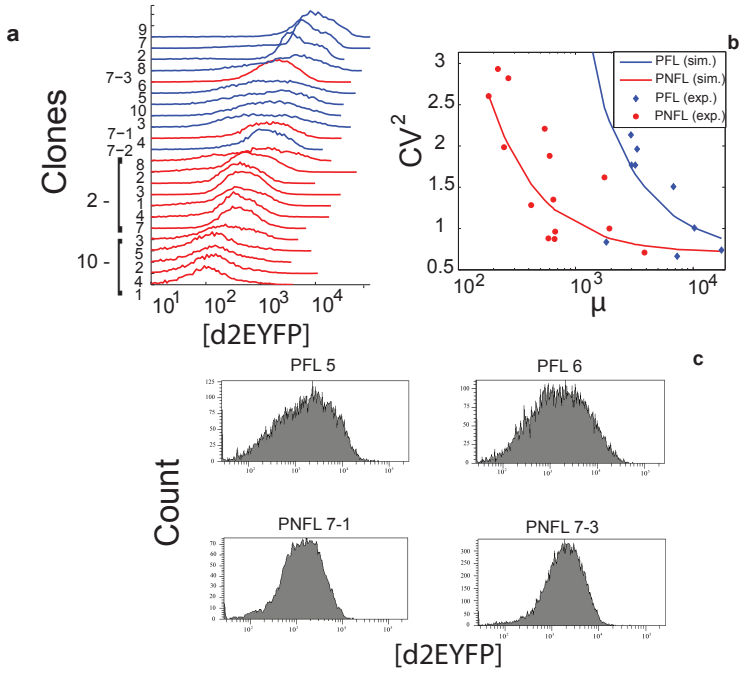


Figure 5.3: **The miRNA-mediated negative feedback loop reduces fluctuations in protein expression. (A)**: FACS-derived histograms of the distribution of d2EYFP fluorescence for all clonal populations of PFL/PNFL cells. **(B)**: Experimental and simulated data for different clones of CV^2 as a function of mean fluorescence. FACS data for all the clonal populations of PFL and PNFL are shown (PFL = blue points, PNFL = red points) ; CV^2 as function of the mean fluorescence has been plotted using the analytical form of CV^2 as described in the first stochastic model for PFL and PNFL (respectively blue and red lines), in order to simulate the experimental data. **(C)** FACS-derived histograms of the distribution of d2EYFP fluorescence in two representative clonal populations of PFL and PNFL cells (PFL 5,6 and PNFL 7-1, 7-3). The reduction of cell-to-cell variability due to the effect of the miRNA is also visible in representative fluorescence microscopy images.

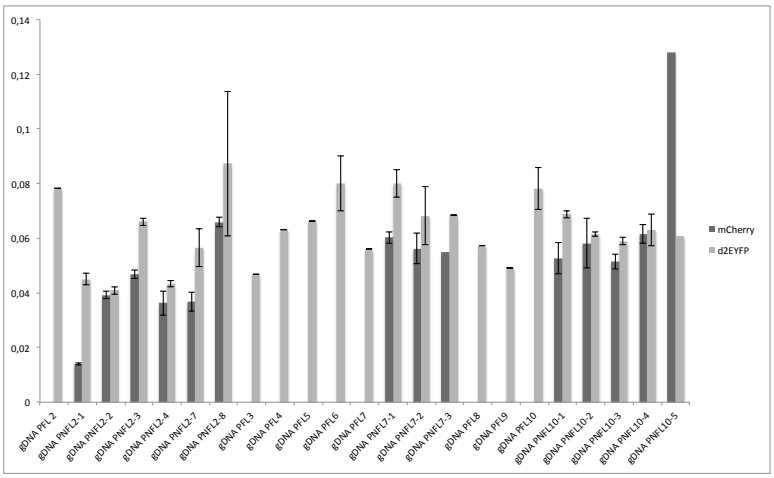


Figure 5.4: **d2EYFP and mCherry integration levels in PFL, and PNFL circuits.** Real-time PCR was performed on genomic DNA extraction from CHO cells to compare the rate of viral integration of each circuit in monoclonal populations.

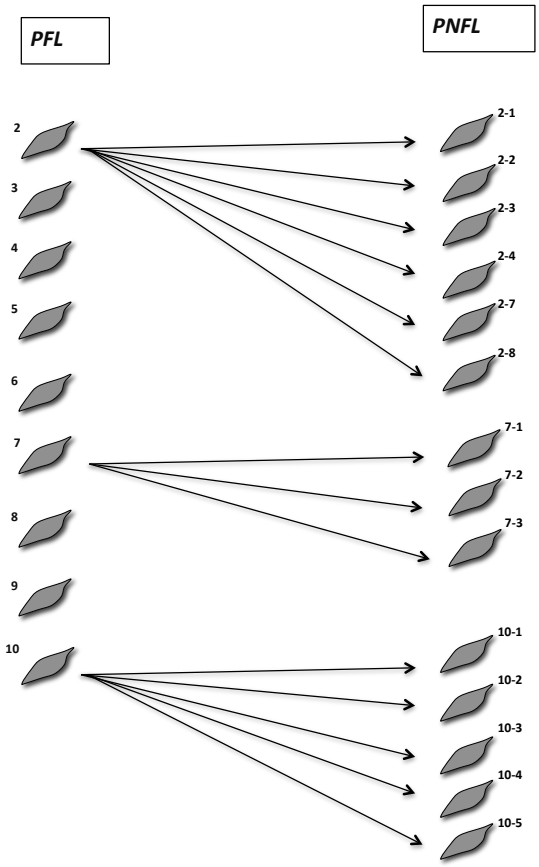


Figure 5.5: **Clone Tree.** PFL clones were derived from the first Lentiviral infection of CHO wild type cells while PNFL clones were derived from the second Lentiviral infection performed on sorted monoclonal populations of PFL cells.

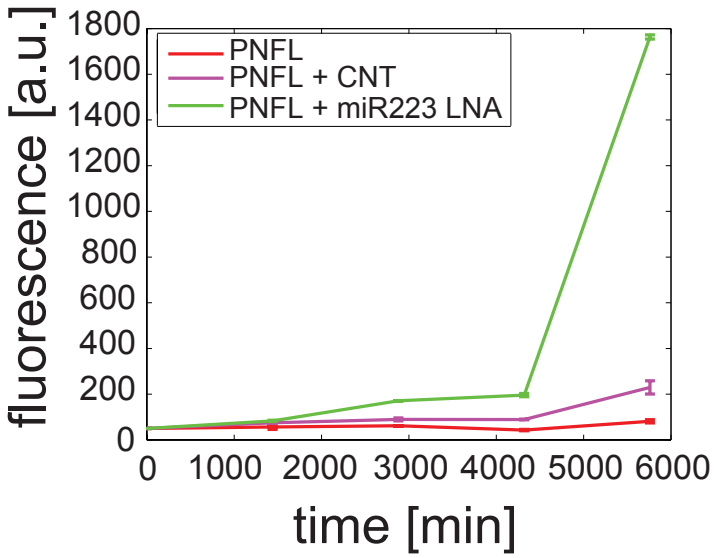


Figure 5.6: **Fluorescence-Activated Cell Sorting (FACS) switch on experiment of PNFL cells transfected with Locked Nucleic Acid 223 (LNA 223).** d2EYFP fluorescence level in PNFL cells trasfected with 150 pmol of LNA 223 (green line), LNA negative control (blue line) and not trasfected (magenta line) were measured at 0 hrs , 24 hrs (1440 min), 48 hrs (2880 min), 96 hrs (5760 min) following removal of Doxycycline at time 0 hrs. Prior to time 0 hrs, PNFL cells were grown in the presence of Doxycycline for 72 hrs and the cells were trasfected after 24 hrs from Doxycycline removal.

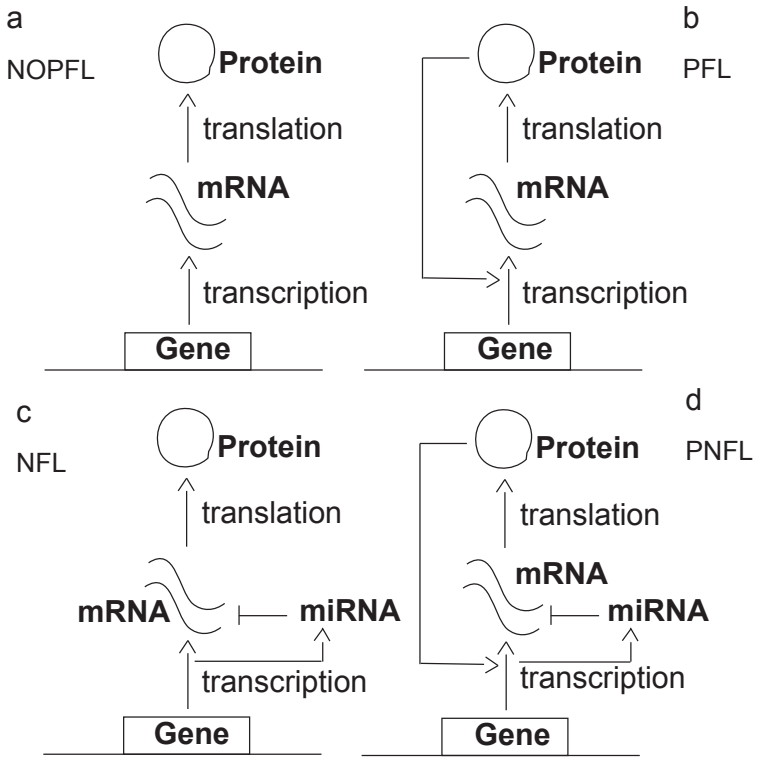


Figure 5.7: **Scheme of NOPFL, PFL, NFL, PNFL circuits. (A)** Open Loop circuit as represented in [128] and corresponding to the NOPFL circuit [126]. (B) PFL where the transcription feedback activates its own transcription. (C) Negative Feedback Loop mediated by a microRNA inducing degradation of the target mRNA. **(D)** A symplified scheme for PNFL system, represented as a combination of PFL and NFL.

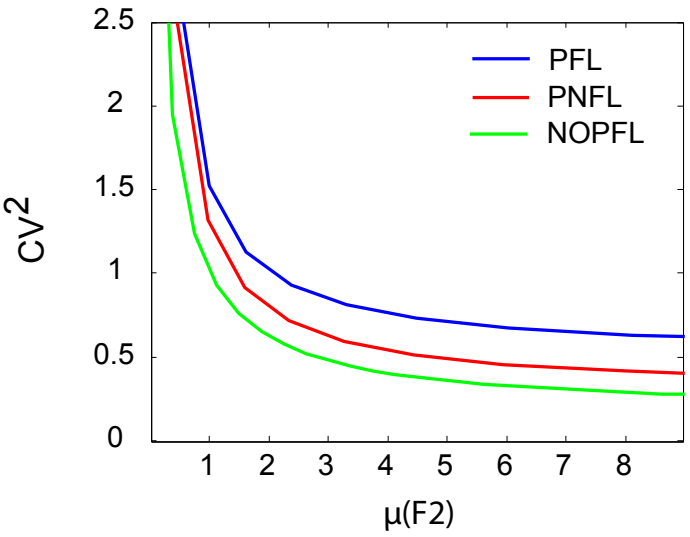


Figure 5.8: CV^2 **as function of the mean fluorescence as derived from the simulation of the Master Equations.** The blue line represents the CV^2 of the PFL while the red line is the CV^2 of the PNFL. The green line is the CV^2 of the NOPFL as preciously estimated in [104].

Chapter 6

The Notch-Delta-Hes1 pathway

Having modeled synthetic circuits as illustrated in Chapters 3, 4, 5, in this chapter I turned my attention to the analysis of endogeneous regulatory networks. The behaviour of endogeneous networks active in single cells can be "modified" as a consequence of the communication with other networks of the same kind expressed in different cells. Indeed, cells are not isolated systems but they react to the external environment, neighbouring cells included. Cell signalling is part of a complex system of communication that governs basic cellular activities and coordinates cell actions. Cells can sense signals generated by intermediate external factors. Once detected the signal causes a chain of biochemical reactions inside the cell itself, usually leading to the actuation of inhibitor of a trascription factor. The set of events is called a signal transduction pathway.

The Notch-Delta pathway is involved in several developmental processes

in which many interactions and genes are implicated: for example, Hes1 is representative for a large family of transcription factors with some specific roles in several developmental processes conserved across the species.

In the following, a biological description of the Notch-Delta-Hes1 pathway is provided.

6.1 Notch Receptor-Delta Ligand Pathway

During embryonic development, communication between cells becomes more and more important: every cell has to correctly differentiate and migrate in order to coordinate its behaviour with the other cells to form every tissue and organ in the organism. The building of an organism from a single cell to a multicellular three-dimensional structure of characteristic shape and size is the result of coordinated gene action that directs the developmental fate of individual cells. So, the evolution and the development of a complex organism is due to the internal features of the single cells, but also to the coordination between individual cells [11]. A single cell can proliferate giving rise to a great number of equal cells that are able to differentiate in other kinds of cells, with different roles depending on the developmental goals: this kind of choice for each cell is the result of an auto-organization process of the group of cells making the initial structure of the organism. One way adopted by the cells to communicate is the physical interaction between proteins expressed on their

membranes: these "membrane proteins" can be distinguished in "receptors", able to transduce the signals represented by the "ligands", present on the membrane of the other cell. The Notch-Delta pathway belongs to this class of contact-based (juxtacrine) signaling pathways. The Notch-Delta signaling pathway is a highly conserved in most multicellular organisms, it is present in all metazoans and it is used to control cell fates through local cell interactions: the Notch receptor is a single-pass transmembrane receptor protein able to bind a set of ligand proteins like Delta/Serrate/Jagged/Dll [24] (depending on the isoform or the particular organism).Once Notch binds one of its ligand present on the surface of another cell, the Notch Intracellular Domain (NICD - the cytoplasmatic part of the Notch protein) is clevead by the enzyme γ-secretase. NICD translocates into the nucleus and cooperates with other DNA-binding proteins to promote transcription of "target genes" (including the bHLH genes of the *Enhancer of split*) with a fast track to the nucleus [23, 24].

Notch signaling is an evolutionary conserved mechanisms: the number of its paralogues differs between species (there are four in mammals, two in *Caenorhabditis elegans*) and one in *Drosophila melanogaster*, but the basic paradigm is common throughout. Thus, Notch signaling is considered to be a general developmental signalling pathway used to direct cell fate and, consequently to build an organism.

Among the biological processes requiring Notch, the best characterized

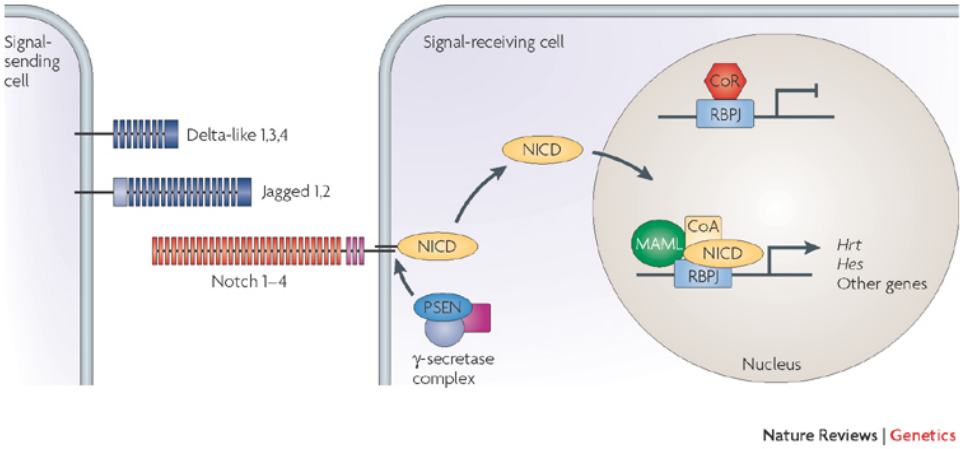


Figure 6.1: **Notch-Delta pathway** When Notch binds the ligand on the membrane of another cell, the Notch Intracellular Domain (NICD) is cleaved and translocates into the nucleus promoting the transcription of its target genes. This "chain of reactions" is the way the two cells can communicate by biochemical interactions. Figure from [63].

ones are: vertebrate somitogenesis, the sharp boundary formation of *Drosophila* wing veins, the "salt and pepper" patterning of *Drosophila eye disk*, the bristle patterning on the fly notum and more [122].

In order to elucidate the mechanisms underlying the Notch-Delta pathway using a mathematical approach, we need first to distinguish the processes considered (see Chapter 7, 8, 9) in two different classes that take into account the dynamical features of biological processes where Notch is involved:

- "Patterning" includes processes in which the genes involved present a

sustained static expression pattern extending over multiple cells (lateral inhibition, stripe and boundary formation, lineage decision);

- "Synchronization" includes processes in which the target genes present a dynamic expression (synchroneus behaviour over time, wave and signal propagation as in the somitogenesis case).

Indeed, depending on the specific process, sometimes Notch-effector genes exhibit a costant level of expression, while other times they have a cyclic expression. Of course, the mechanisms behind these different processes can be sometimes interlinked and each of this regulations appears in different contexts between different kinds of cells.

6.2 Role of Notch-Delta in Patterning

Since the Notch-Delta signaling is involved in different developmental processes, several genes are regulated by the Notch receptor. One of this processes is known as *lateral inhibition* [4], [11] where the Notch receptor is able to amplificate and make stronger differences between communicating cells. Lateral inhibition is an example of "patterning": cells communicating in a tissue present a nonuniform stable expression of the Notch receptor and target genes.During the lateral inhibition process, the cells approach a state of non uniformity in which a group of cells presents a high level of expression of Notch, while a second group presents a high level of Delta: this is the

case of the *Drosophila* disc eye formation, where a single cells differentiate in photoreceptor neurons at uniformly spaced positions, surrounded by other photoreceptor cells.

The mechanism underlying the lateral inhibition process is the emergence of a sort of self-organization from which a "chosen" cell is selected: this cell transmit the signals to its neighbours to differentiate into other types of cells, giving rise to a "salt and pepper" pattern. This contact-dependent signaling leads to a "competition" with only one cell emerging as a "winner", strongly expressing Delta and preventing its neighbours from doing the same. Lateral inhibition seems to be essential also in the case of neuron development [119], in order to avoid the formation of too many neuronal cells in the epidermal tissue, which is lethal. Lateral inhibition has been studied also from a mathematical point of view: in [130] and [131], the authors derive a model of lateral inhibition able to regulate the "salt and pepper" pattern formation starting from a uniform tissue of equal cells.

From a mathematical point of view, "patterning" includes not only the "salt and pepper" pattern typical of lateral inhibition, but all those cases in which cells do not present a uniform state although they are identical. Thus, the formation of stripes and the boundary formation (as in the case of Drosophila wings) are included in the lateral inhibition process [26].

During the proneural differentiation, another kind of behaviour is observed, in which the target genes present an oscillatory behaviour in some cells, while reaching a steady state level in other cells. In [67], [123] and [119] this behaviour is described: in particular, the expression of some genes involved in the mantainence of the undifferentiated state are seen to oscillate in expression during this phase, while, when the tissue is coming to a stable structure, the cells which will be differentiated into neurons, stop this cyclic expression and switch to a sustained expression of the same genes.

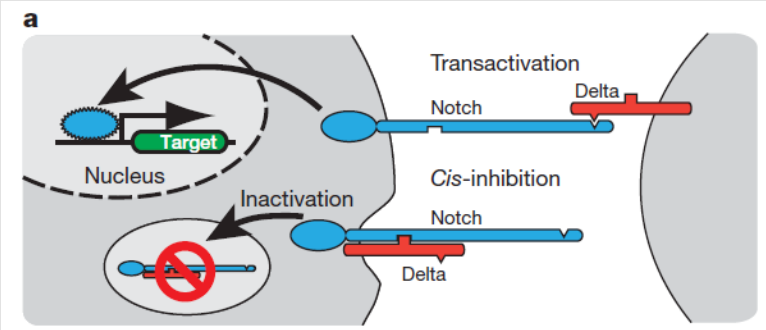


Figure 6.2: **Transactivation and cis-inhibition.** Notch receptor can bind to Delta ligand available on the membrane of another cell (transactivation) or it can decrease the amount of free Delta in the same cell directly attaching to it (cis-inhibition). Figure from [131].

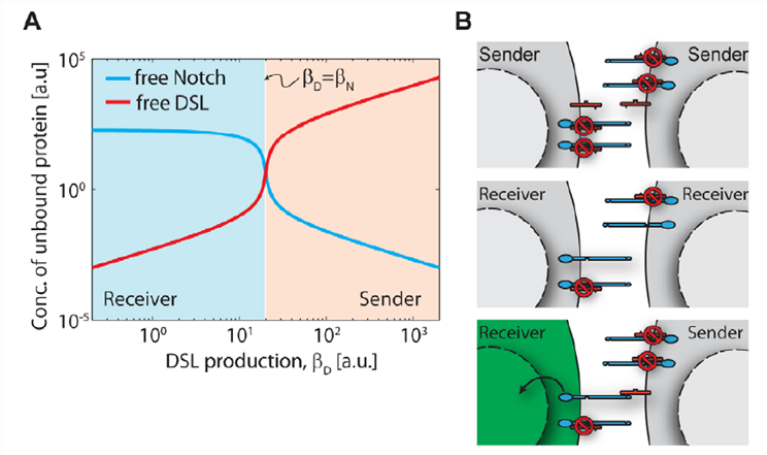


Figure 6.3: **Receiver and Sender cells.** Two neighbouring cells can be both "sender"' or "receiver"' or one of the first type and the other of the second one. A cell with a low amount of free Notch and a high amount of free Delta is defined to be a sender, otherwise a receiver. Figure from [130].

6.3 Role of Notch-Delta in Synchronization: the Hes1 ultradian oscillator

Qualitatively different from lateral inhibition is the case of somitogenesis. Somitogenesis subdivides the paraxial mesoderm of vertebrate embryos into a series of homologous subunits, the somites: they form sequentially at the anterior end of the presomitic mesoderm (PSM) and in vertebrates, they give rise to skeletal muscles, cartilage, tendons, endothelial cells, and dermis. A molecular oscillator known as the "segmentation clock" directs cyclic

expression of genes in the PSM and is coupled with the progression of the somitogenesis. Notch signaling is clearly central to somitogenesis as it can be seen by mutating components of Notch pathway, which causes somites to be eventually formed but segment boundaries are irregular and randomly positioned [106], [43]. To model the process of somite formation, the *Clock and Wavefront model* (Figure 6.4) has been proposed as a possible explanation of this phenomenon ([113], [41], [116], [96], [14]). This model is characterized by three fundamental components: an autonomous oscillator in each cell (the clock), a global synchronization of the clock of the population, a gradient of morphogens able to determine the point of the embryo in which the somite will be formed (the wavefront) and thus determining the onset of the segmentation program. In details, the cells involved form the Presomitic Mesoderm (PSM) exhibit synchronized autonomous oscillations thanks to the cell-cell coupling mediated by the Notch-Delta pathway; these cells are able to propagate a transcriptional wave starting from the posterior part of the embryo; this wave slows down thanks to an external gradient of FGF and Wnt until it stops propagating at the anterior part of the embryo, thus determining the exact point in which a new couple of somites will form. Somites are formed in couples along the vertical axis of the embryo and starting from its anterior part while it continues to elongate at the posterior. Each somites is thus determined by the period of the oscillations able to modify their number and size: somites also presents a rostral-caudal polarity. The biological factors

taking part in the process of somite boundary formation and in the determination of the rostral-caudal polarity are not completely determined and their mechanism is still controversial and under investigation: in [43], [106], [54] it has been proven that somite boundaries are formed independently of the clock, as the molecular oscillator is needed to time segment formation and to confer the correct rostro-caudal polarity. The mammalian *Hairy*-related transcription factors code for negative regulators that are able to regulate their own transcription (Hes) display an oscillatory expression pattern and are key genes for proper somite formation. Genes belonging to the family of the

Oscillations of the mouse genes Hes7 and Hes1 have been observed in the mouse PSM [71], and Hes1 oscillations were also observed in fibroblast cell cultures synchronized by serum shock ([64]). The mechanism underlying the cyclic gene expression of Hes1 (or its homologous corresponding to the species) is a negative delayed autoregulatory feedback loop, as shown in Figure 8.2. During somitogenesis, Hes genes have periods of 90 min, 120 min, 30 min, respectively in chick, mouse, zebrafish embryo.

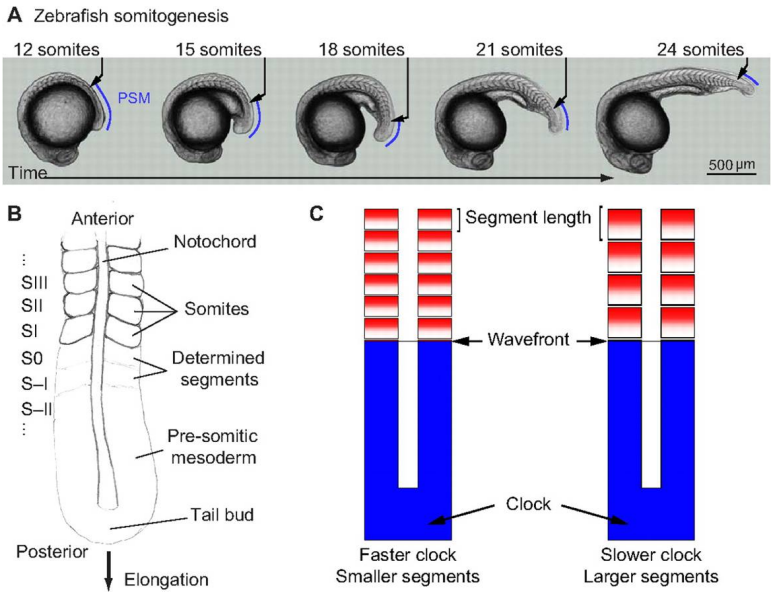


Figure 6.4: **Somitogenesis: Clock-Wavefront model.** (A) Advancing stages of zebrafish embryo somitogenesis, lateral view. Somites bud sequentially from the posterior unsegmented tissue: the pre-somitic mesoderm (PSM, blue line). Embryonic growth and the segmentation process are coordinated. (B) Dorsal view of the posterior of a segmenting vertebrate embryo. Somites form in pairs on each side of the notochord.(C) The Clock and Wavefront mechanism of vertebrate segmentation. The Clock is created by oscillating gene expression in the PSM (blue) and the Wavefront by a posteriorly moving front that arrests the oscillations of the Clock. Resulting segments (red/white) have a length dependent on the period of the Clock and the velocity of the Wavefront. Figure from [100].

Chapter 7

Modeling the Notch-Delta pathway: *Patterning*

In Chapter 6, I provided a biological description of the Notch-Delta pathway, of its features and the processes it is involved in. In this chapter, biological cellular mechanisms that bring to a stable inhomogeneous configuration in a tissue (i.e. "patterning") are studied and analysed from a mathematical point of view. A review of models proposed in literature is provided and a new model on the system composed by Receptor-Reporter-Ligand is proposed, able to emulate different biological processes.

7.1 Modeling cell-to-cell signalling and pattern formation

Cell-to-cell signalling across a tissue can be modeled using the network framework. In this case, each cell is a node and the cell-to-cell signaling (mediated by Notch-Delta pathway) is represented as an edge between two nodes.
From a mathematical point of view, several analytical and numerical strategies have been developed to study a network of cells or a lattice: often, PDEs are useful to describe the diffusion of a substance able to create a communnication between cells or from an external signal like a morphogen.

Coupling strength and the direction of the interaction are essential components to reveal the dynamics of signal propagation across cells. Models of the dynamical properties of Notch pathway across a cell population have been proposed in [89] and in [131], [130], [122].

In [131] the authors use mathematics and experiments to model the cell-to-cell signaling mediated by Notch-Delta. The authours distinguish trans-interactions occurring between Notch on one cell and Delta on the other cell, from cis-interactions occurring between Notch and Delta on the same cell. This cis-interaction induces a mutual inactivation between the receptor and the ligand in the same cell which is able to generate an ultrasensitive switch between mutually exclusive sending (high Delta/low Notch) and receiving

(high Notch/low Delta) signalling states, facilitating the formation of sharp boundaries and lateral inhibition patterns. Indeed the authours were able to simulate a non-uniform distribution of sender and receiver cells, giving rise to a "salt and pepper" pattern starting from an homogeneous population of cells.

Figure 6.2 (from [131]) shows the cleavage of NICD due to the transactivation between Notch and Delta and the inhibition of the same molecules when interacting in the same cell; in Figure 6.3 (from [131]) the mechanism regulating the behaviour of a cell is presented: increasing the production of Delta on the membrane, the cell present an increasing amount of ligand on the membrane and a decreasing amount of Notch, thus becoming a "sender cell", otherwise, when the production of Delta is less than the production of Notch, the cell becomes a "receiver cell".

7.2 Modeling Notch - Delta in a single cell

I considered as a starting point two models first described in [130]: both models are constituted by an equation for free Delta, an equation for free Notch and one equation for the reporter. The first model, shown in Figure 7.2, is the simple Lateral Inhibition model (LI) in which the NICD can activate the transcription of the reporter that inhibits Delta transcription thus diminishing the amount of free ligand on the membrane, able to bind the receptor of the other cells (trans-activation); the second model, in Figure 7.3, is the Lat-

eral Inhibition Mutual Inactivation model (LIMI) in which the same system is considered with the inclusion of a mutual inactivation term between the receptor and the ligand in the same cell (cis-inhibition).

The ODEs describing the LI model are:

$$\frac{dN}{dt} = \beta_N - d_N N - \frac{D_{ext} N}{K_t} \tag{7.1}$$

$$\frac{dD}{dt} = \beta_D \frac{1}{1+R^2} - \frac{N_{ext} D}{K_t} - d_D D \tag{7.2}$$

$$\frac{dR}{dt} = \beta_R \left(\frac{(N D_{ext})^n}{k_{RS} + (N D_{ext})^n} \right) - d_R R, \tag{7.3}$$

in the case of the LIMI model the equation become:

$$\frac{dN}{dt} = \beta_N - d_N N - \frac{D_{ext} N}{K_t} - \frac{ND}{K_c} \tag{7.4}$$

$$\frac{dD}{dt} = \beta_D \frac{1}{1+R^2} - \frac{N_{ext} D}{K_t} - d_D D - \frac{ND}{K_c} \tag{7.5}$$

$$\frac{dR}{dt} = \beta_R \left(\frac{(N D_{ext})^n}{k_{RS} + (N D_{ext})^n} \right) - d_R R. \tag{7.6}$$

The two models are identical but for the cis-inactivation between Notch and Delta in the same cell given by the term $-\frac{ND}{K_c}$ in Eqs. 7.4 and 7.5; N and D represent the Notch and Delta amounts on the membrane with the extra equation for the reporter R.

Referring to both models, Notch in Eq.7.4 is considered to be produced at a constant rate β_D accounting for the leakiness of its promoter, with a linear degradation term (d_N as degradation rate). The Notch production

rate decreases also because of binding with both external and internal Delta. The Delta production rate in Eqs. 7.2, 7.5 is the sum of reporter protein itself by the binding with the external and internal Notch (N_{ext} and N) and by a linear degradation (d_D).

The LI and LIMI models can be adimensionalised using the following variable changes:

$$N = \frac{N}{N_0}, D = \frac{D}{D_0}, R = \frac{R}{R_0}, t = \gamma_R t,$$

where $N_0 = D_0 = \gamma k_t$ and $R_0 = k_{DR}$. Thus the LIMI model becomes

$$\begin{aligned}
\tau \frac{dN}{dt} &= b_N - N - D_{ext} N - \frac{ND}{k_c} && (7.7)\\
\tau \frac{dD}{dt} &= b_D \frac{1}{1+R^2} - N_{ext} D - D - \frac{ND}{k_c} && (7.8)\\
\tau \frac{dR}{dt} &= b_R \left(\frac{(N\, D_{ext})^n}{K_{RS} + (N\, D_{ext})^n} \right) - R; && (7.9)
\end{aligned}$$

choosing, $N_0 = D_0 = \gamma K_t = 1$, $R_0 = K_{DR} = 1$, $\tau = \frac{\gamma_R}{\gamma} = 1$, $b_N = \frac{\beta_N}{\gamma N_0}$, $b_D = \frac{\beta_D}{\gamma D_0}$, $b_R = \frac{\beta_R}{\gamma_R R_0}$, $k_c = \frac{K_c}{K_t}$, $K_{RS} = \frac{k_{RS}\gamma_s K_t}{N_0 D_0}$. Similar equations can be derived for the LI model.

In addition to the LI and LIMI model, I considered also the MI (mutual inactivation) model described in Figure 7.1. This model can be described by the following adimensional equations:

$$\tau \frac{dN}{dt} = b_N - N - D_{ext} N - \frac{ND}{k_c} \tag{7.10}$$

$$\tau \frac{dD}{dt} = b_D \frac{1}{1+R^2} - N_{ext} D - D - \frac{ND}{k_c}. \tag{7.11}$$

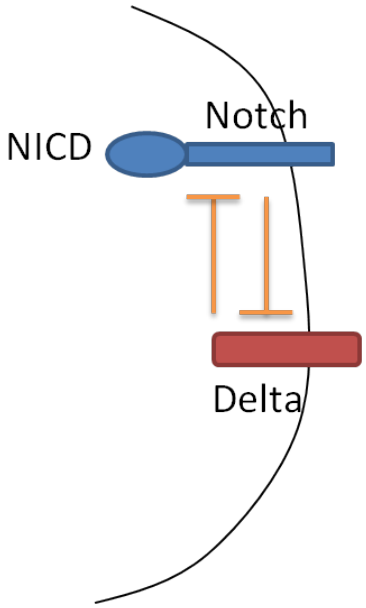


Figure 7.1: **MI model.** Mutual Inactivation model: Notch and Delta can bind in the same cell inhibiting each other.

It is possible to show that, for a single cell, all the three models (LI, LIMI, MI) can support either a sender or a receiver behaviour, i.e. they can express high Notch and low Delta and viceversa, depending on the values of b_D and b_N, as shown in Figs. 7.4, 7.5 and 7.6; when two interacting cells are considered, the bifurcation analysis in Figure 7.7 shows that the MI model

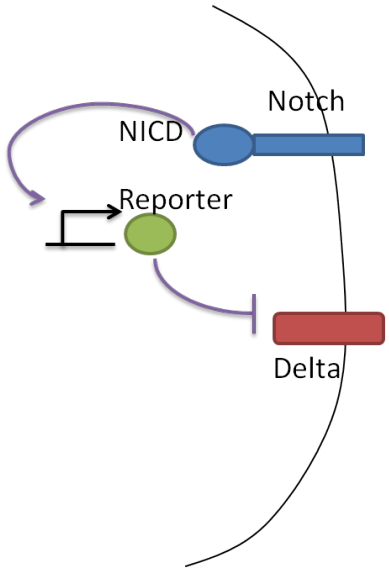


Figure 7.2: **LI model.** Lateral Inhibition model: in the same cell, only the activation of Notch on the reporter is considered and the reporter on the Delta.

cannot support a sender/receiver behaviour, i.e. the two cells can either be both senders or both receivers, but it is never possible to have one sender cell and one receiver cell. This implies that to obtain an "asimmetric" behaviour, it is necessary to have an extra regulatory intermediate interaction as in the LI and LIMI case. Bifurcation diagrams for the LI and LIMI models describing two cells scenario are reported in Figs. 7.8 and 7.9. These diagrams show that both LI and LIMI models can approach different equilibria depending on the starting point including a sender/receiver configuration.

Parameter	Value	**Meaning**
b_N	10	constant production for free Notch
d_N	1	degradation rate for free Notch
b_D	100	constant production for free Delta
d_D	1	degradation for free Delta
b_R	24300	maximal transcription rate for the reporter gene
d_R	1	degradation rate for the reporter gene
k_c	1	mutual inactivation strength
k_t	1	trans-activation strength
K_{RS}	300000	Hill constant for NICD and reporter binding site
n	3	Hill coefficient for NICD and reporter binding site

Table 7.1: Table of parameters for the MI, LI, LIMI proposed models

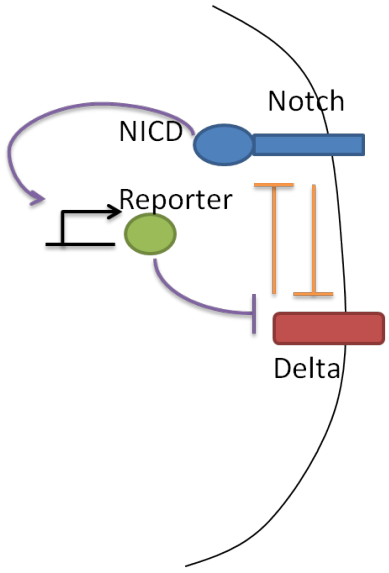


Figure 7.3: **LIMI model.** Lateral Inhibition and Mutual Inactivation model: in the same cell, both interactions with the promoter and between Notch and Delta are taken into account.

7.3 Modeling Notch - Delta in a system of $\mathcal{N}$ cells

In order to consider a network of cells, I assumed a lattice in which each cell in the network can communicate with the other six neighbouring cells as shown in Figure 7.10 ([32]).

The LIMI model presented in Section 7.2 for one cell, can be extended to $\mathcal{N}$ cells and rewritten as:

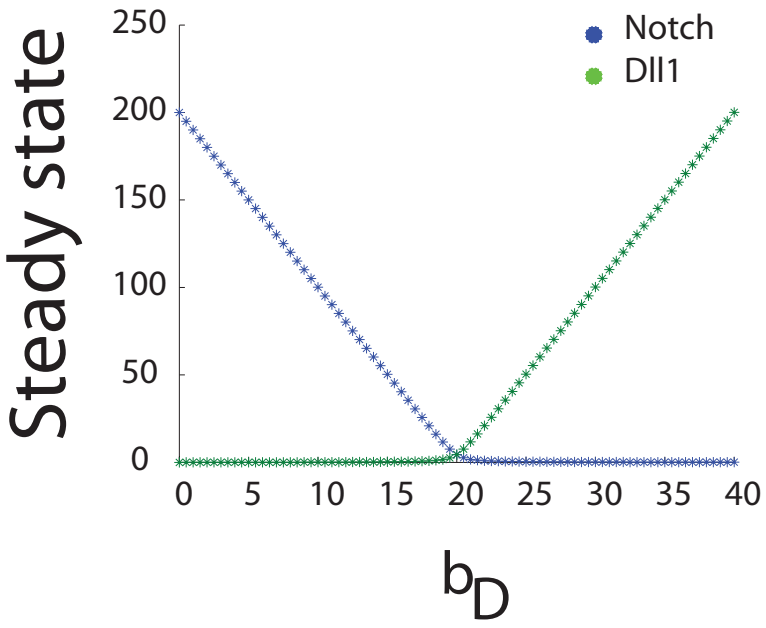


Figure 7.4: **Sender-Receiver behaviour for a cell with MI system.** Bifurcation diagram for a single cell with respect to the parameter b_D (i.e. production of free Delta). Notch (in blue) and Delta (in green) are shown at the steady state. When the production term is low Notch level is higher than Delta level, thus the cell being a receiver; otherwise the cell become a sender.

$$\frac{dN_i}{dt} = b_N - N_i - \sum_{j=]i[} D_j N_i - N_i D_i \tag{7.12}$$

$$\frac{dD_i}{dt} = b_D \frac{1}{1+{R_i}^2} - \sum_{j=]i[} N_j D_i - D_i - N_i D_i \tag{7.13}$$

$$\frac{dR_i}{dt} = b_R \left(\frac{(N_i \sum_{j=]i[} D_j)^n}{K_{RS} + (N_i \sum_{j=]i[} D_j)^n} \right) - R_i. \tag{7.14}$$

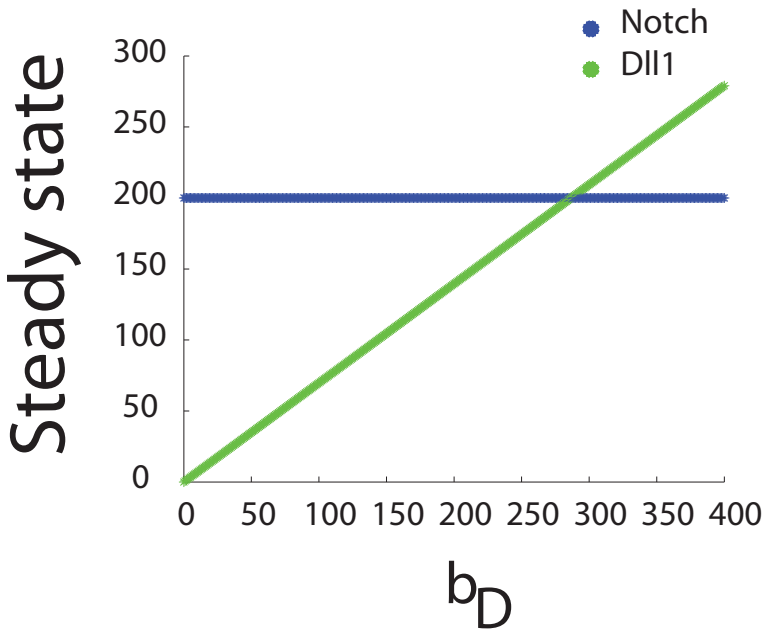


Figure 7.5: **Sender-Receiver behaviour for a cell with LI system.** Bifurcation diagram for a single cell with respect to the parameter b_D (i.e. production of free Delta). Notch (in blue) and Delta (in green) are shown at the steady state. When the production term is low Notch level is higher than Delta level, thus the cell being a receiver; otherwise the cell become a sender.

with symbol $]i[$ we indicate the set of the indexes defining the six cells neighbouring of the cell i. The term $\sum_{j=]i[} D_j N_i$ has the same meaning of the component $-\frac{D_{ext}N}{K_t}$ for the LIMI model of one cell, thus obtaining a system of $3 \times \mathcal{N}$ equations.

In order to check if there exists a parameters' region where patterning can occur, one needs to identify, conditions for the divergence from the homogeneous

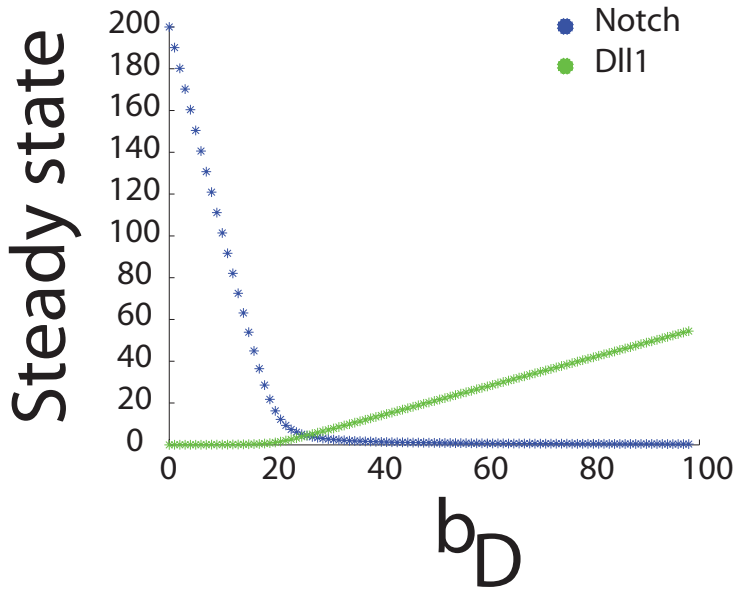


Figure 7.6: **Sender-Receiver behaviour for a cell with LIMI system.** Bifurcation diagram for a single cell with respect to the parameter b_D (i.e. production of free Delta). Notch (in blue) and Delta (in green) are shown at the steady state. When the production term is low Notch level is higher than Delta level, thus the cell being a receiver; otherwise the cell become a sender.

steady-state solution , i.e. when all the cells reach the same equilibrium (no sender/receiver pattern). To this end I performed a linear stability analysis around the homogeneuos steady state [111], using the Maximum Lyapunov Exponent (MLE). Equations 7.12-7.14 for the uniform steady state (when all the cells are equal) become:

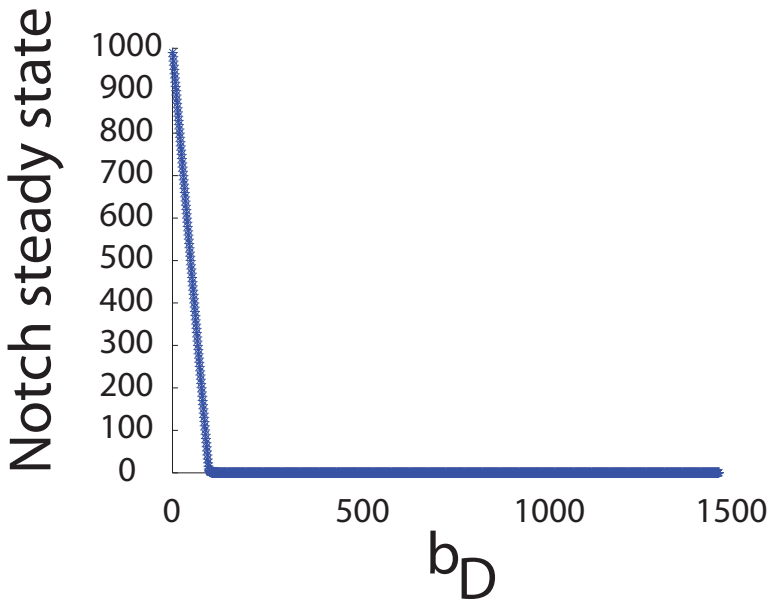


Figure 7.7: **Bifurcation diagram for a two-cell MI system.** Bifurcation diagram representing the Notch state of the first state depending on the parameter b_D (free Delta production). The system can be monostable and the stable equilibrium is a situation of receiver/receiver or sender/sender, but never of sender/receiver.

$$b_N - N - DN - ND = 0 \tag{7.15}$$

$$b_D \frac{1}{1 + {p_2}^2} - ND - D - ND = 0 \tag{7.16}$$

$$b_R \left(\frac{(ND)^n}{K_{RS} + (ND)^n} \right) \left(\alpha + (1 - \alpha) \frac{k{p_0}^2}{{p_0}^2 + {p_2}^2} \right) - R = 0. \tag{7.17}$$

To simplify the analysis we express the Jacobian as the sum of two ten-

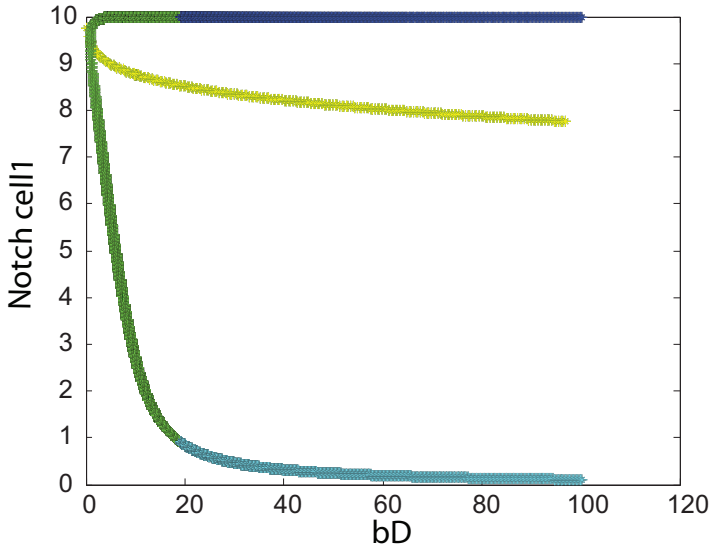


Figure 7.8: **Bifurcation diagram for a two-cell LI system.** Bifurcation diagram representing the Notch state of the first state dependng on the parameter b_D (free Delta production). The system can be monostable or multistable; depending on the color, the possible equilibria are receiver/receiver (green or yellow) and sender/receiver (cyan and blue); no sender/sender state is reachable.

sor products of matrices, one for the internal dynamics and the other for interactions with the other cells [105]:

$$J = I_{\mathcal{N}} \otimes H + M \otimes B,$$

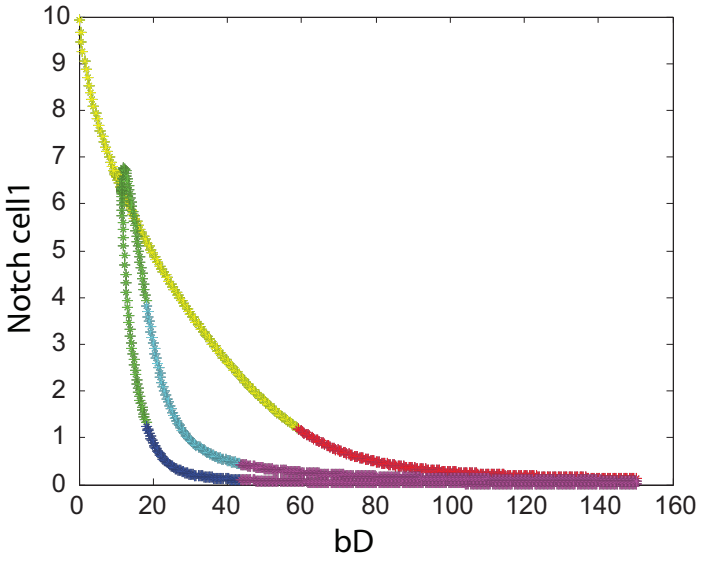


Figure 7.9: **Bifurcation diagram for a two-cell LIMI system.** Bifurcation diagram representing the Notch state of the first state dependng on the parameter b_D (free Delta production). The system can be monostable or multistable; depending on the color, the possible equilibria are receiver/receiver (green or yellow), sender/sender (red or magenta) and sender/receiver (cyan and blue).

where $\otimes$ is the tensor product as defined for matrices A, B of dimension k as

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1k}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \cdots & a_{kk}B \end{pmatrix},$$

$I_{\mathcal{N}}$ represents the $\mathcal{N} \times \mathcal{N}$ identity matrix, $\mathcal{N}$ is the number of cells involved

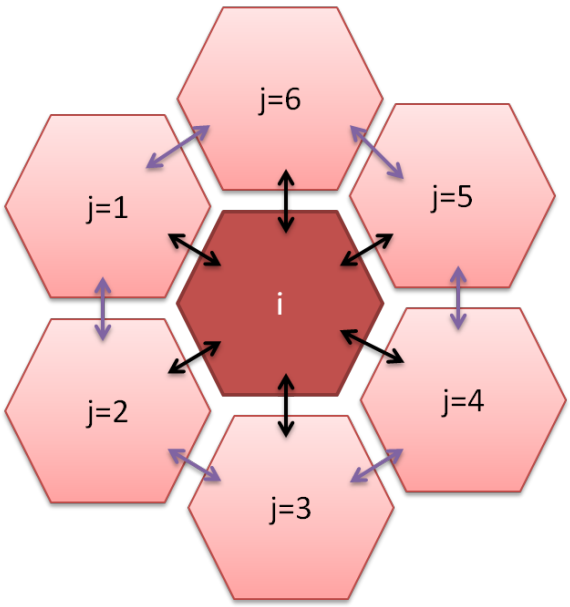


Figure 7.10: **Scheme of the cell-to-cell interactions.** Cells are depicted as hexagons; the cell i communicates with the six neighbouring cells (black arrows). Also the surrounding cells can interact each other following the scheme of the purple arrows.

in the interactions in question, $H_{i,j} = \frac{\partial \dot{q}_i}{\partial q_j}$ is the change in production of species i for a change in species j in the same cell, while the change in production of the species i for a change in species j in a neighboring cell is given by $B_{ij} = \frac{\partial \dot{q}_i}{\partial \langle q_j \rangle}$. Then matrix M is the connectivity matrix, $M_{ij} = \frac{1}{6}$ if cell i is a neighbour of cell j, 0 otherwise. Once the Jacobian has been simplified, its eigenvalues can be rewritten as the eigenvalues of the matrices $H + q_k B$, where q_k are the eigenvalues of the connectivity matrix M. In our

case, as in [105], $q_k \geq -0.5$, so the MLE can be computed only considering $q_k = -0.5$. So, we report H and B for both the LI and LIMI models:

$$H_{LI} = \begin{pmatrix} -(1+D) & 0 & 0 \\ 0 & -(1+N) & -\beta_D \frac{mR^{m-1}}{(1+R^m)^2} \\ \beta_R n K_{RS} \frac{D(ND)^{n-1}}{(K_{RS}+(ND)^n)^2} & 0 & -1 \end{pmatrix};$$

$$B_{LI} = \begin{pmatrix} 0 & -N & 0 \\ -D & 0 & 0 \\ 0 & \beta_R n K_{RS} \frac{N(ND)^{n-1}}{(K_{RS}+(ND)^n)^2} & 0 \end{pmatrix};$$

and

$$H_{LIMI} = \begin{pmatrix} -(1+D+\frac{D}{K_c}) & -\frac{N}{K_c} & 0 \\ -\frac{D}{K_c} & -(1+N+\frac{N}{K_c}) & -\beta_D \frac{mR^{m-1}}{(1+R^m)^2} \\ \beta_R n K_{RS} \frac{D(ND)^{n-1}}{(K_{RS}+(ND)^n)^2} & 0 & -1 \end{pmatrix};$$

$$B_{LIMI} = \begin{pmatrix} 0 & -N & 0 \\ -D & 0 & 0 \\ 0 & \beta_R n K_{RS} \frac{N(ND)^{n-1}}{(K_{RS}+(ND)^n)^2} & 0 \end{pmatrix}.$$

Following [130], I plotted the MLE as function of b_N and b_D, for two different values of the Hill coefficient $n = 1$ and $n = 3$. Results are shown in Figs. 7.11. When MLE is negative, the cell population is homogeneous, whereas when MLE becomes positive, the homogeneous solution is unstable hence patterning may emerge.

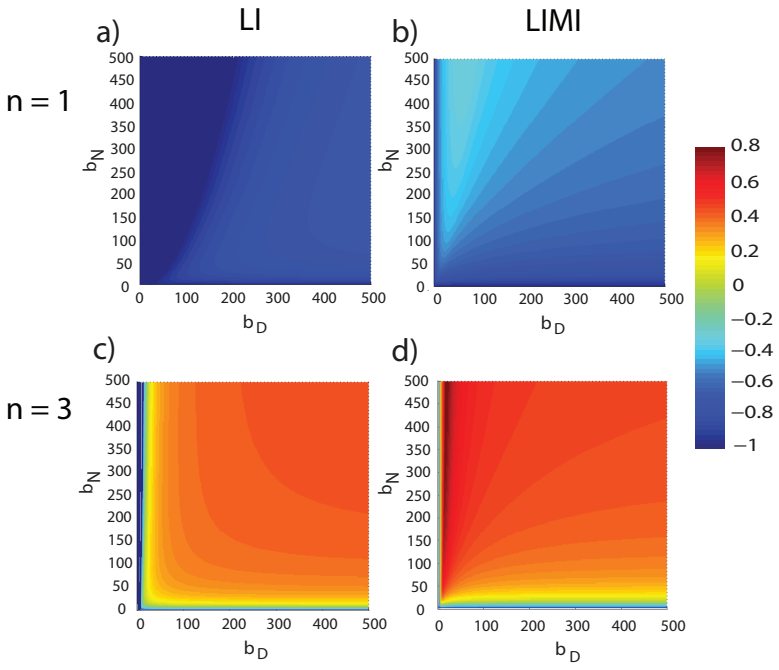


Figure 7.11: **Master Lyapnov Exponent.** MLE for (a) LI with $n = 1$ (b) LIMI with $n = 1$ (c) LI with $n = 3$ (d) LIMI $n = 3$. When $n = 1$ no pattern can be obtained, while when $n = 3$, patterns are more easily obtained.

By comparing the results in Figure 7.11 for the LI and LIMI model obtained, we find the same results as in [130], that the LIMI is characterized by a higher capacity (with respect to the LI model) of creating patterns also in the case of a low cooperativity between the NICD and the promoter of the gene reporter (case $n = 1$).

I also visualized the patterns emerging in the population for different values of the parameters (b_N, b_D) within the "patterning" parameters' region.

To this end, I considered a grid of 20×20 cells, and derived a connectivity matrix L for the network with periodic boundary conditions. I chose random initial conditions and let the LIMI system of ODEs i 7.7-7.9 evolves to non-steady state. At each time point, the contributions deriving from all the other cells can be computed as

$$\underline{v} = \begin{pmatrix} \frac{1}{k_1} \sum_{j=]1[} \underline{x}_j \\ \vdots \\ \frac{1}{k_{\mathcal{N}}} \sum_{j=]\mathcal{N}[} \underline{x}_j \end{pmatrix},$$

which can be rewritten as

$$\underline{v} = -(L \otimes I_{\mathcal{N}})\underline{x} + (K \otimes I_{\mathcal{N}})\underline{x},$$

where the vector $\underline{x} = (\underline{x}_1, ..., \underline{x}_{\mathcal{N}})'$ with

$$\underline{x}_i = \begin{pmatrix} N_i \\ D_i \\ R_i \end{pmatrix}.$$

Matlab codes for performing this simulations are reported in Appendix A. Figure 7.12 shows some patterns which can be obtained for different choices of (b_N, b_D), such as "salt and pepper", "stripes" and "boundary formation". Similar results can be obtained for the LI model.

7.4 Conclusions

In conclusion, we have shown that a single model composed by Notch-Delta pathway and a reporter is able to reproduce stable patterns in a lattice of cells. In the following chapters, these results will be compared to the case in which the target gene presents an oscillatory expression, leading to the formation of other phenomena difficult to obtain in the case of genes with a sustained expression.

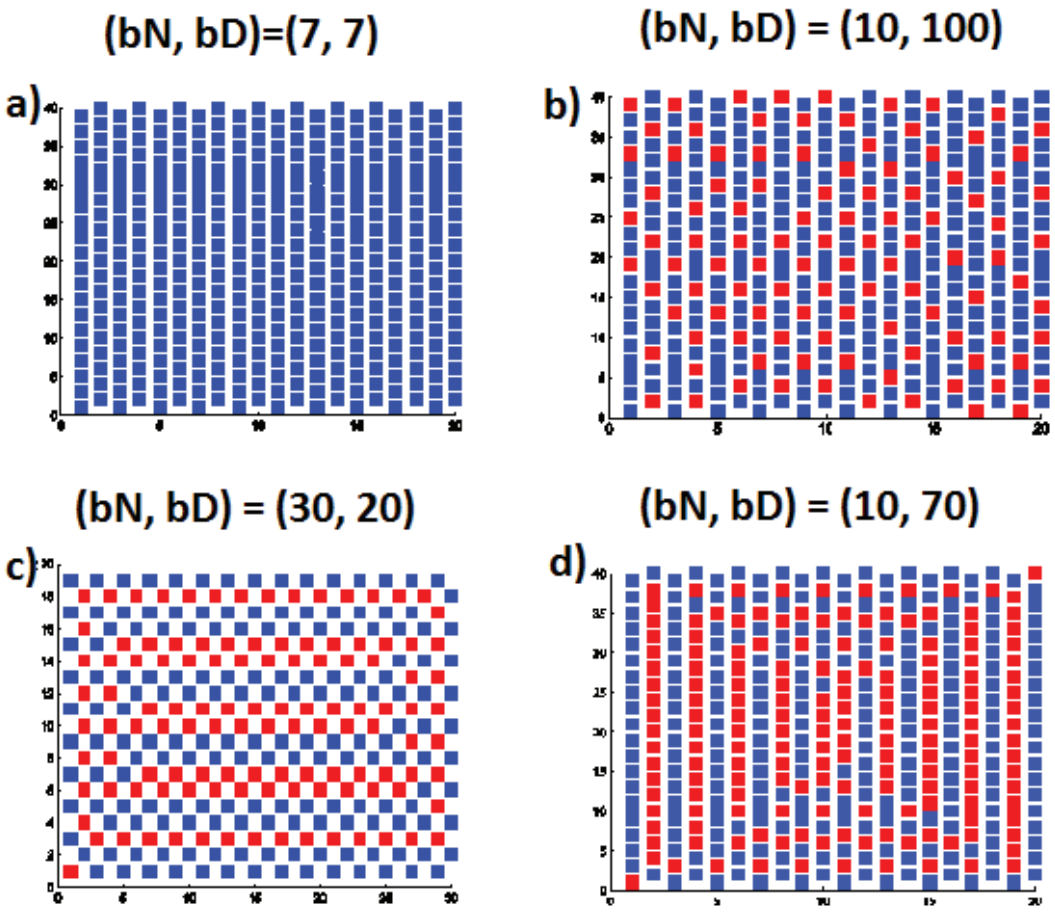


Figure 7.12: **Patterning.** Simulated tissue of communicating cells simulated using Eqs. 7.7-7.9 for the LIMI model: each cell influences the behaviour of the 6 neighbouring cells. (a) uniform state (b) salt an pepper (c) boundary formation (d) stripes. Each pattern has been obtained with a different choice of parameters (b_D, b_N).

Chapter 8

Modeling the Notch-Delta pathway: *Synchronization*

Here I developed a model to take into account the case in which the Notch receptor regulates the expression of a reporter gene which exhibits an autonomous oscillatory expression. The reporters gene considered is representative of the family of Hes/Her genes involved in several developmental processes (refer to Chapter 6): after introducing some ideas on synchronization, models from literature are reviewed for the autonomous oscillator, then a new model is proposed and the collective behaviour of the coupled oscillating cell population is studied.

8.1 Introduction to *Synchronization*

First recognized in 1665 by Christiaan Huygens, synchronization phenomena are abundant in science, nature, engineering and social life. He discovered that a couple of pendulum clocks hanging from a common support had synchronized, i.e., their oscillations coincided perfectly and the pendula moved always in opposite directions: systems as diverse as clocks, singing crickets, cardiac pacemakers, firing neurons and applauding audiences exhibit a tendency to operate in synchrony. These phenomena are universal and can be understood within a common framework based on modern nonlinear dynamics [110]. Physiological rhythms, for example, are central to life: some rhythms are mantained throughout life and even a brief interruption leads to death. The rhythms interact each other and with the external environment: variation of rhytms outside the normal limits, or appearence of new rhytms where none existed previously, is associated with a disease, [55].

Synchronization can be defined as an adjustment of dynamics of the individual oscillators due to their interaction which eventually converge to the same trajectory. Synchronization of oscillators is considered a spontaneous and collective phenomen arising from nearly identical autonomous oscillators coupled to each other. Beside what is called "synchronization", another process is considered called "entrainment", in which an external signal stimulates "disconnected units" leading them to a common behaviour: in this case, neither coupling between the different components nor autonomous oscilla-

tors are necessary. In the case of synchronization, oscillators starting from different phases converge to the same trajectory depending on the intrinsic properties of the oscillators considered and on the coupling features, such as topology, coupling strength and mode (by diffusive effects or large/short range of action). In the case of Huygens' pendulum clocks, each pendulum is an oscillator that can modify the behaviour of the other one thanks to the information of its movement transmitted by the beam (their communication channel). As a consequence, in order to develop an analysis on the synchronization features of our system, a model for the autonomous oscillator is necessary as the wiring between the other oscillating cells too.

8.2 A review of dynamical models of the Hes oscillator

The Hes1 transcription factor is expressed at cyclic intervals in specific tissues and cell types, as described in Chapter 6, Section 6.3. Several models have been proposed in literature to describe the autonomous mechanism of the "Hes1 oscillator": most of these models use the Delayed Differential Equations formalism ([80], [95], [65]), other are based on Stochastic models ([15], [31], [148]) or by Ordinary Differential Equations [64], [56]. The most common model found in literature is the one described in [80]: the mechanism of her1/her7 oscillator gene (member of the Hes family of TFs) during the

zebrafish segmentation process is described by a system of two DDEs (one for the mRNA and the other for the protein):

$$\frac{dx_1}{dt} = \frac{k}{1+\frac{{x_2}^2(t-\tau)}{{p_0}^2}} - cx_1 \tag{8.1}$$

$$\frac{dx_2}{dt} = ax_1(t-T_p) - bx_2. \tag{8.2}$$

x_1 and x_2 represents the mRNA and the protein concentrations respectively: the x_1 transcription rate is given by a decreasing function of the protein with a delay equal to τ. This term represents the inhibiting action of the protein on the mRNA transcription and a linear degradation term. The x_2 protein translation rate is proportional to the mRNA concentration with a delay T_p and to a linear degradation term. In [80], also a numerical study is proposed on the role of Notch-Delta in the somite formation in *Zebrafish*: specifically, a single *her1* gene oscillator described by Eqs. 8.1 and 8.2 is shown to synchronize across cells via the Notch-Delta pathway. Notch is supposed to be proportional to the quantity of Delta in the other cell, thus contributing to the production of *her1* gene.

The model described by 8.1 and 8.2 was also used in [65] to show that Hes7 is crucial for the somitogenesis organization also in mouse. The authors show that oscillations with a period of 120 minutes compatible with what has been observed in mouse, can be generated by setting a delay $\tau = 37$ minutes.

Note that the period of oscillations in mouse is of 120 minutes, longer then in *Zebrafish*, where the period is 20 minutes.

Another model of autonomous oscillation of Hes1 is described in [64], this time using the ODE formalism:

$$\frac{dx}{dt} = -Axz + By - Cx \tag{8.3}$$

$$\frac{dy}{dt} = -Dy + \frac{E}{1+x^2} \tag{8.4}$$

$$\frac{dz}{dt} = -Axz + \frac{F}{1+x^2} - Gz, \tag{8.5}$$

where x is the protein concentration, y is the mRNA concentration and z is a Hes-interacting factor. In this model, the oscillations are due to the presence of an external factor z: when the equation corresponding to z is removed, the system is not able to sustain oscillations.

A more complex model of Hes1 was presented in [94]: here, the Hes1 oscillator is modeled using a system of 5 DDEs which takes into account also the nucleo-cytoplasmic transport, Hes dimerisation, and differential stability of Hes monomers and dimers. The authors show that extending the model by incorporating periodic forcing of the Hes1 circuit by cyclic phosphorylation of the protein Stat3, time delays and differential stability act synergistically to generate large amplitude oscillatory solutions that match.

Another mathematical model in which Notch-Delta pathway is involved is reported in [103], where somite segmentation clock is investigated. Simulations were performed using a ODE model of six equations involving the

Hes7 oscillator, *Lunatic Fringe* gene (Lfng), Notch and Dll1 as coupling between cells: authors demonstrated that Notch signaling is responsible for the synchronization of the autonomous oscillators and that Lfng represses Notch activity in neighbouring cells by modulating Dll1 function. This work proposes Lfng as an extra component acting between Hes7 and Dll1, and interfering with Notch1 as shown in Figure 8.1.

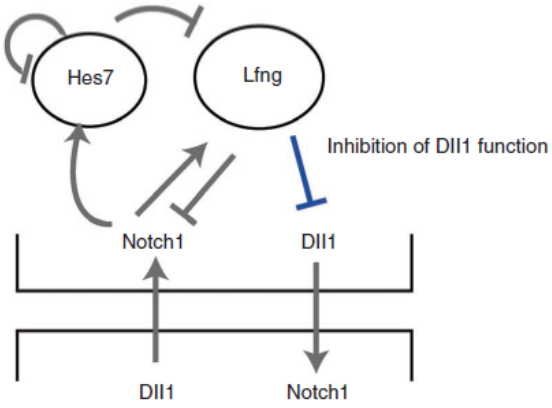


Figure 8.1: **Hes 7 network as proposed in [103].** Lunatic Fringe (Lfng) is considered as intermediate step regulating the intercellular coupling of Hes7 by inhibiting Dll1 function in addition to Notch1 activity.

8.3 A new model for the Hes1 molecular oscillator

In this section I derived a simple deterministic mathematical model for the Hes molecular oscillator in mouse depicted in Figure 8.2. I first reduced the DDE model described in Eqs. 8.1 and 8.2, to a monodimensional model, described by 8.6) and I then transformed this model to a system of ODEs described by Eqs. 8.7-8.9, so that the explicit time delay can be removed from the equations.

The three dimensional model I derived belongs to the class of Goodwin oscillators [86].

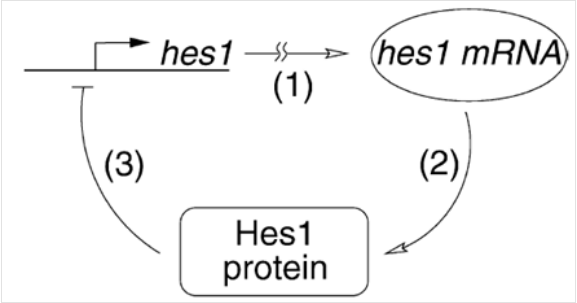


Figure 8.2: **Internal autonomous oscillator as in [86].** The delayed negative feedback loop is proposed as the core mechanism leading to the Hes gene oscillations.

8.3.1 A DDE monodimensional model

I considered a simple, reduced model based on one delayed differential equation with the unique delay on the negative feedback loop of Hes1 on itself. The obtained model is described by the following equation:

$$\frac{dx_1}{dt} = \frac{k}{1 + \frac{{x_1}^2(t-\tau)}{{p_0}^2}} - cx_1. \tag{8.6}$$

I simplified the usual model (Eqs. 8.1-8.2) by considering a single equation for the Hes1 mRNA and protein. Choosing $\tau = 55$ minutes I could obtain the experimentally observed period of 120 minutes.

In Figs. 8.3, 8.4, 8.5, I performed a series of bifurcation analyses describing the changes in the amplitude of the oscillation for different sets of the parameters. Figures 8.6, 8.7, 8.8 show the bifurcation analyses related to the period of oscillations.

8.3.2 A Goodwin oscillator model

It is known from literature that a delayed autoregulatory negative feedback is the most simple case of an oscillating system [114], [101]: this is known as the Goodwin oscillator [57]. This kind of system has been modeled either with one or more delayed equations (as in the case of the Hes7 model in [80]) or by removing the dependence on the explicit time delay, by considering a three dimensional ODE system of the following two kinds:

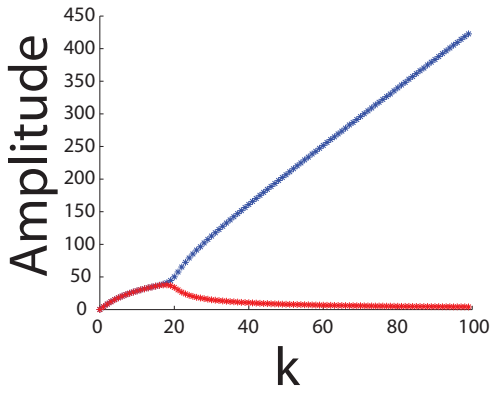


Figure 8.3: **Bifurcation analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter k.** Parameter k represents the maximal transcription rate for the mRNA repression: the higher is the value, the wider are the amplitude of the oscillator.

(A) a simple model with an high level of cooperativity between the repressor and the responding promoter;

(B) a model with a low cooperativity but with a nonlinear term for the degradation.

Here, the three equations represent the Hes1 mRNA and the Hes1 protein in the unfolded and folded form.

The simplest form of type (A) model is described by the following equations:

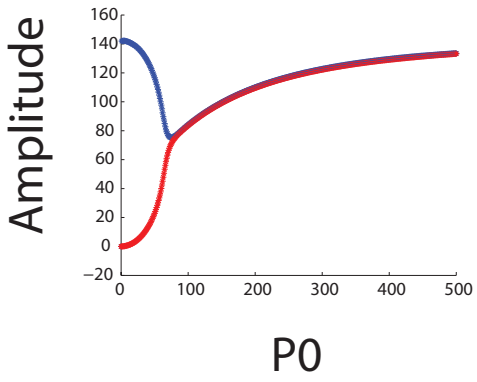


Figure 8.4: **Bifurcation analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter p_0.** Parameter p_0 represents the strength of the inhibition of the protein (x_3) on the mRNA (x_1): oscillations are lost for very high values of p_0.

$$\frac{dx_1}{dt} = b_H\left(\alpha + (1-\alpha)\frac{{p_0}^n}{{p_0}^n + {x_3}^n}\right) - d_1 x_1 \tag{8.7}$$

$$\frac{dx_2}{dt} = v_1 x_1 - (d_2 + K_f)x_2 \tag{8.8}$$

$$\frac{dx_3}{dt} = K_f x_2 - d_2 x_3 \tag{8.9}$$

where

- x_1, x_2, x_3 are respectively the mRNA, unfolded protein, folded protein concentrations,
- parameter α represents the leakiness of the promoter,

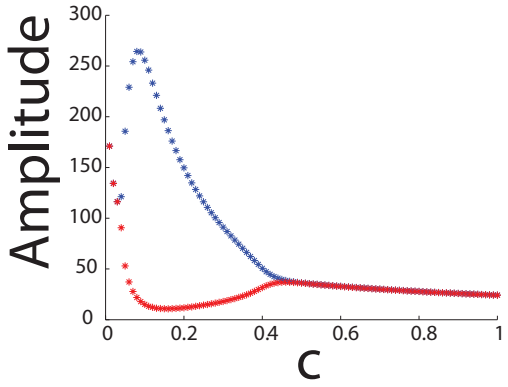


Figure 8.5: **Bifurcation analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter c.** Parameter c represents the degradation rate of the mRNA: an oscillatory behaviour is observed for the system for low values of degradation.

- d_1 and d_2 the degradation rates,
- K_f the folding rate for the protein,
- b_H and v_1 the maximal transcription rates for the mRNA and the protein,
- p_0 and n the coefficients of the Hill autorepressing function of Hes1.

The model can present oscillations if $n > 8$.

For the type (B), the equations are:

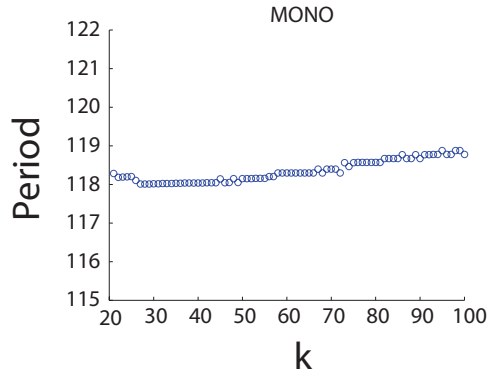


Figure 8.6: **Period analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter k.** Parameter k does not affect the period of the oscillation.

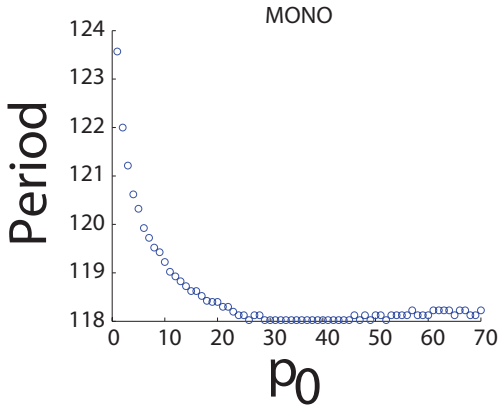


Figure 8.7: **Period analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter p_0.** Decreasing the parameter, the period decreases too.

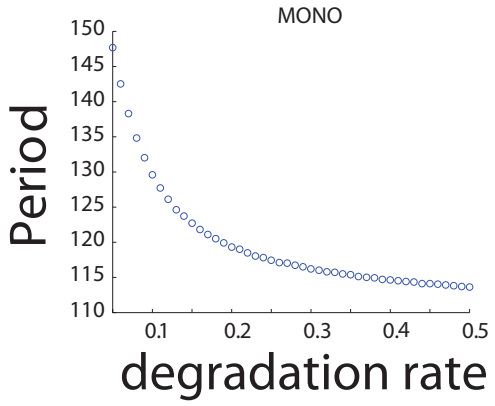


Figure 8.8: **Period analysis for the mono-dimensional DDE model of Hes1, with respect to the parameter** c**.** Decreasing the parameter, the period decreases too.

$$\frac{dx_1}{dt} = \tau\left(b_H\left(\alpha + (1-\alpha)\frac{k{p_0}^2}{{p_0}^2 + {x_3}^2}\right) - d_H x_1\right) \tag{8.10}$$

$$\frac{dx_2}{dt} = \tau\left(v_1 x_1 - (d_2 + K_f)x_2\right) \tag{8.11}$$

$$\frac{dx_3}{dt} = \tau\left(K_f x_2 - d_2\left(\frac{x_3}{K_p + x_3}\right)\right) \tag{8.12}$$

where x_1, x_2, x_3 represent respectively the Hes1 mRNA, the unfolded and folded form of the protein; the parameters are the same as model (A), with τ a time rescaling constant and K is the Hill constant for the nonlinear degradation: this Hill function with exponent equal to 1 is sufficient to generate oscillations.

Since the high rate of cooperativity, represented by coefficient n, necessary to reach a sustained oscillation for the model (A) is a very strong hypothesis to accept biologically, I preferred to use model (B) with a nonlinear protein degration rate. Thus, with the parameters reported in Table 8.3.2, model (B) presents stable oscillations (Figure 8.9). The adimensional constant τ can be used to change the time scale and hence the period of the oscillations. When $\tau = 0.037$ the model presents a 120 minutes oscillation period; as during mouse somitogenesis, with $\tau = 0.055$ the new period is 90 min as in chick somitogenesis, while with $\tau = 0.15$ the period becomes 30 min as in zebrafish somitogenesis.

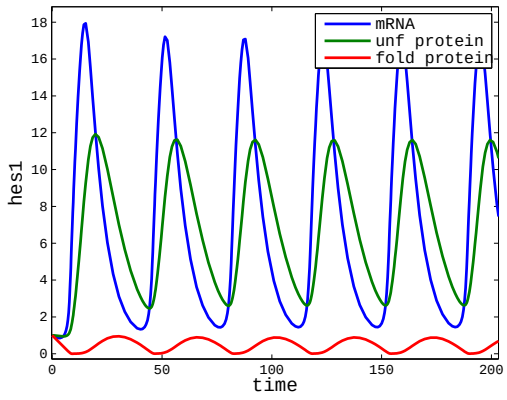


Figure 8.9: **Simulation for the three-dimensional model of Hes1.** Blue line for the mRNA oscillations; green line for the unfolded protein oscillation; red line for the folded protein oscillations.

Parameter	Value	**Meaning**
k	33	maximal transcription rate
τ	37	delay of Hes1 inhibition
p_0	0.1	Hill constant for mRNA transcription pf Hes1
α	0.5	partition term between NICD and Hes1 protein effect on Hes1 mRNA
b_H	1	maximal transcription rate for Hes1 promoter
v_1	1	maximal transcription rate for Hes1 unfolded protein
K_p	0.01	constant Hill for Hes1 protein nonlinear degradation
d_H	1	degradation rate for Hes1 mRNA
d_2	1	degradation rate for Hes1 protein
K_f	0.135	protein folding rate

Table 8.1: Table of parameters for the Goodwin oscillator

In Figs. 8.10, 8.11, 8.12 I perform a bifurcation analyses illustrating the change in amplitude of the oscillations for different sets of parameters. In Figs. 8.13, 8.14, 8.15 bifurcation plots related to the period of the oscillations are shown.

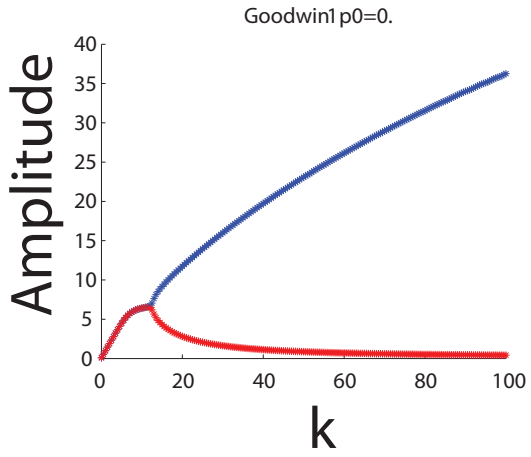


Figure 8.10: **Bifurcation analysis for the three-dimensional model of Hes1, with respect to the parameter** k**.** Parameter k represents the maximal transcription rate for the mRNA repression: the higher is the value, the wider are the amplitude.

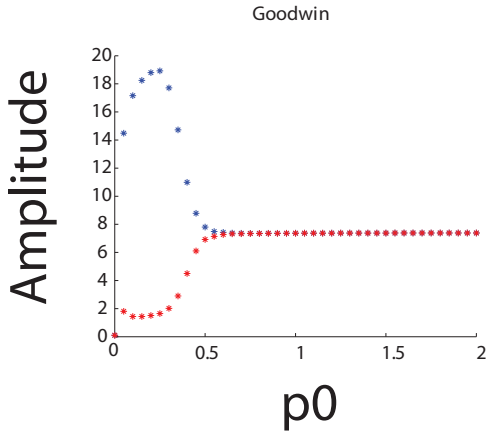


Figure 8.11: **Bifurcation analysis for the three-dimensional model of Hes1, with respect to the parameter p_0.** Parameter p_0 represents the Hill constant for the repression of the protein on the mRNA.

8.4 An ODE model of the Notch-Delta-Hes pathway

In what follows, I describe the derivation of a model of the Hes molecular oscillator interacting with the Notch-Delta pathway. In order to do this, I used the LI and LIMI models described in Chapter 7, Eqs. 7.12-7.14, substituting the simple "reporter" with the "repressor" which represents a generic gene of the Hes/Her family and it is responsible of the delayed negative feedback loop inducing oscillations. The LI and LIMI models become the LI-Repressor and LIMI-Repressor systems (LIR and LIMIR), as shown in Figs. 8.16.

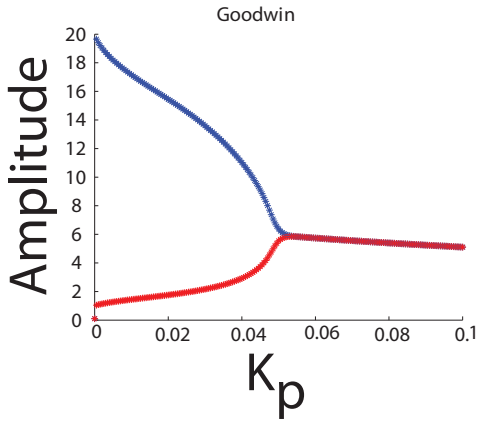


Figure 8.12: **Bifurcation analysis for the three-dimensional model of Hes1, with respect to the parameter K_p.** Parameter K_p represents the Hill constant for the nonlinear degradation term for the protein: the lower is the value, the wider are the amplitude.

The LIR model is described by the following system of ODEs:

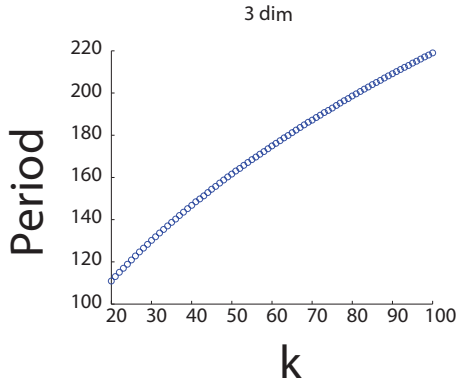


Figure 8.13: **Period analysis for the three-dimensional model of Hes1, with respect to the parameter k.** The period increases with parameter k.

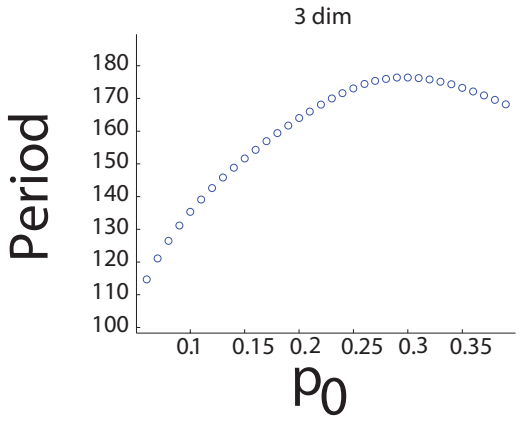


Figure 8.14: **Period analysis for the three-dimensional model of Hes1, with respect to the parameter p_0.** The period increases with significative values of parameter p_0.

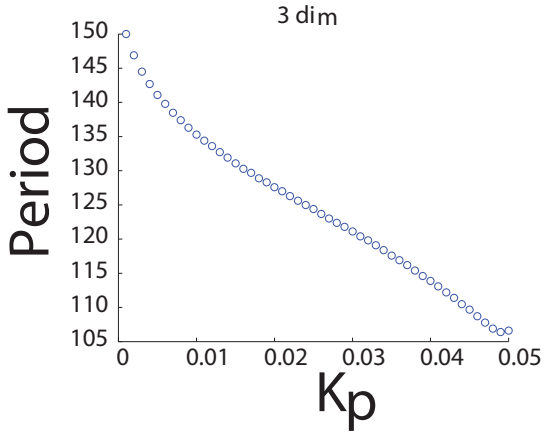


Figure 8.15: **Period analysis for the three-dimensional model of Hes1, with respect to the parameter K_p.** The period decreases when parameter K_p increases.

$$\frac{dN}{dt} = \tau(b_N - d_N N - \frac{D_{ext} N}{K_t}) \tag{8.13}$$

$$\frac{dD}{dt} = \tau(b_D \frac{1}{1 + {p_f}^2} - d_D D - \frac{N_{ext} D}{K_t}) \tag{8.14}$$

$$\frac{dR}{dt} = \tau(b_R \left(\frac{(N D_{ext})^n}{K_{RS} + (N D_{ext})^n} \right) \left(\alpha + (1 - \alpha) \frac{k {p_0}^2}{{p_0}^2 + {p_f}^2} \right) - d_R R) \tag{8.15}$$

$$\frac{dp}{dt} = \tau(b_p R - (d_p + K_f) p) \tag{8.16}$$

$$\frac{dp_f}{dt} = \tau \left(K_f p - d_f \frac{p_f}{K_p + p_f} \right). \tag{8.17}$$

$$\tag{8.18}$$

The LIMIR is described by the following system of ODE:

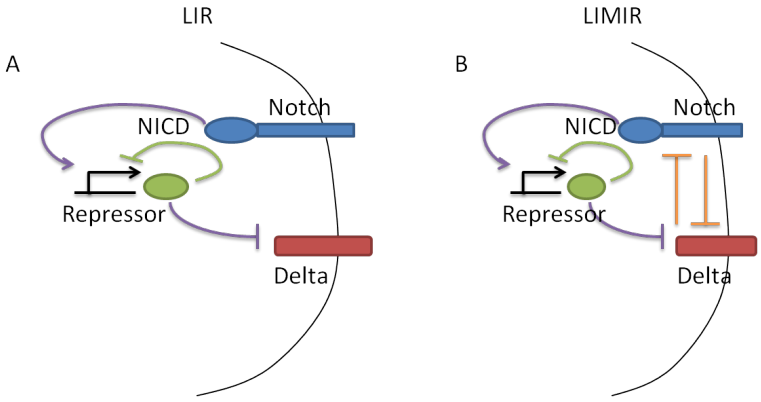


Figure 8.16: **LIR and LIMIR models.** (A) Lateral Inhibition Repressor model (LIR) and (B) Lateral Inhibition Mutual Inactivation Repressor model (LIMIR).

$$\frac{dN}{dt} = \tau(b_N - d_N N - \frac{D_{ext}N}{K_t} - \frac{ND}{K_c}) \tag{8.19}$$

$$\frac{dD}{dt} = \tau(b_D \frac{1}{1+{p_f}^2} - \frac{ND}{K_c} - d_D D - \frac{N_{ext}D}{K_t}) \tag{8.20}$$

$$\frac{dR}{dt} = \tau(b_R \left(\frac{(ND_{ext})^n}{K_{RS}+(ND_{ext})^n}\right)\left(\alpha + (1-\alpha)\frac{k{p_0}^2}{{p_0}^2+{p_f}^2}\right) - d_R R) \tag{8.21}$$

$$\frac{dp}{dt} = \tau(b_p R - (d_p + K_f)p) \tag{8.22}$$

$$\frac{dp_f}{dt} = \tau\left(K_f p - d_f \frac{p_f}{K_p + p_f}\right). \tag{8.23}$$

$$\tag{8.24}$$

The LIR and LIMIR models consist of 5 ODEs, two for free Notch and Delta on the membrane and the other three describing the core mechanism

of the repressor component with the Goodwin model formulated in Section 8.3.2. The two models differ for the term $-\frac{ND}{K_c}$, indicating the cis-inactivation between Notch and Delta hyphotised in [131]. The parameter τ is useful to modify the period of the oscillator, thus the model can be adapted to different situations. Parameters values are reported in Table 8.4: two different cooperativity rates are considered between Notch and the repressor promoter ($n = 1$ and $n = 3$).

In Figs. 8.17 and 8.18 the behaviour of LIR and LIMIR models open loop with $n = 1$ is reported. As expected, for the LIR system, Notch is at a constant level, while the repressor's oscillation leads to an oscillatory behaviour of the ligand Delta too; on the contrary, Notch is seen to oscillate in the LIMIR model since its oscillation is guided by the mutual inactivation term. For the LIR system a simple analytical analysis can be perfomed, as in the following.

The adimensionalized models for LIR and LIMIR can be obtained by setting $N = \frac{N}{N_0}, D = \frac{D}{D_0}, R = \frac{R}{R_0}, t = \gamma_R t$, where $N_0 = D_0 = \gamma k_t$ and $R_0 = k_{DR}$. If one considers a set of $\mathcal{N}$ coupled cells, the adimensionalised LIR model becomes:

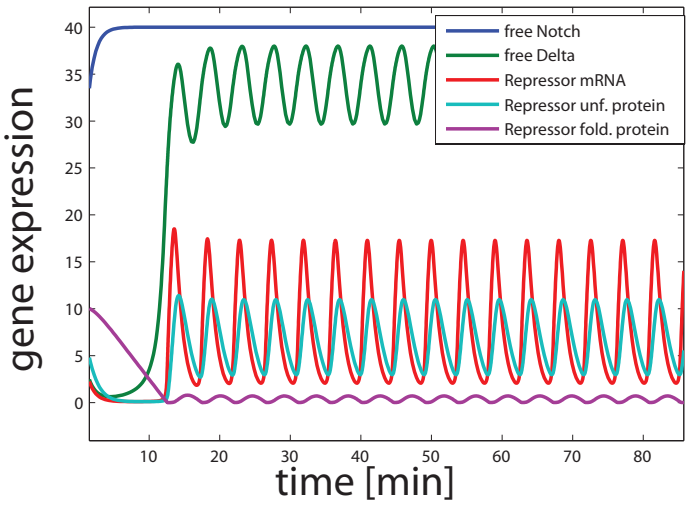


Figure 8.17: **Oscillations in a single cell LIR system.** Numerical simulation to discover the dynamical behaviour of the LIR system with $n = 1$.

$$\frac{dN_i}{dt} = \tau(b_N - N_i - \sum_{j=]i[} D_j N_i) \tag{8.25}$$

$$\frac{dD_i}{dt} = \tau(b_D \frac{1}{1 + {p_{2i}}^2} - \sum_{j=]i[} N_j D_i - D_i) \tag{8.26}$$

$$\frac{dR_i}{dt} = \tau(b_R \left(\frac{(N_i \sum_{j=]i[} D_j)^n}{K_{RS} + (N_i \sum_{j=]i[} D_j)^n} \right) \left(\alpha + (1 - \alpha) \frac{k {p_0}^2}{{p_0}^2 + {p_{2i}}^2} \right) - R_i) \tag{8.27}$$

$$\frac{dp_{1i}}{dt} = \tau(b_p R_i - (1 + K_f) p_{1i}) \tag{8.28}$$

$$\frac{dp_{2i}}{dt} = \tau \left(K_f p_{1i} - \frac{p_{2i}}{K_p + p_{2i}} \right); \tag{8.29}$$

$$\tag{8.30}$$

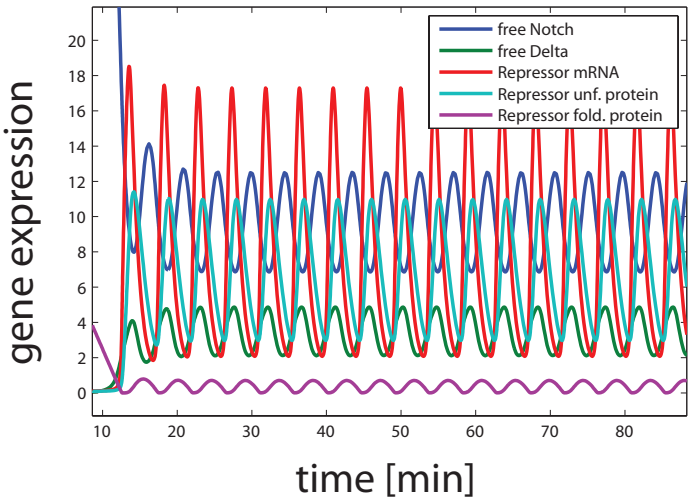


Figure 8.18: **Oscillations in a single cell LIMIR system.** Numerical simulation to discover the dynamical behaviour of the LIMIR system with $n = 1$.

whereas the adimensional LIMIR model is

$$\frac{dN_i}{dt} = \tau(b_N - N_i - \sum_{j=]i[} D_j N_i - N_i D_i) \tag{8.31}$$

$$\frac{dD_i}{dt} = \tau(b_D \frac{1}{1+{p_{2i}}^2} - \sum_{j=]i[} N_j D_i - D_i - N_i D_i) \tag{8.32}$$

$$\frac{dR_i}{dt} = \tau(b_H + b_R \left(\frac{(N_i \sum_{j=]i[} D_j)^n}{K_{RS} + (N_i \sum_{j=]i[} D_j)^n} \right) \left(\alpha + (1-\alpha) \frac{k{p_0}^2}{{p_0}^2 + {p_{2i}}^2} \right) - R_i) \tag{8.33}$$

$$\frac{dp_{1i}}{dt} = \tau(b_p R_i - (1+K_f) p_{1i}) \tag{8.34}$$

$$\frac{dp_{2i}}{dt} = \tau \left(K_f p_{1i} - \frac{p_{2i}}{K_p + p_{2i}} \right). \tag{8.35}$$

$$\tag{8.36}$$

8.5 Analytical study of the LIR and LIMIR models

To proceed with an analytical study of these models, I first reduced both models to only two equations corresponding to Notch and Delta by replacing the oscillator equations with an explicit periodic forcing function, obtaining for the LIR model the following equations:

$$\frac{dN_i}{dt} = b_N - N_i - \epsilon \sum_{j=]i[} D_j N_i \tag{8.37}$$

$$\frac{dD_i}{dt} = b_D f(t, \sum_{j=]i[} D_j N_i) - \epsilon \sum_{j=]i[} N_j D_i - D_i. \tag{8.38}$$

Parameter	Value	**Meaning**
k	33	maximal transcription rate
τ	37	delay of Hes1 inhibition
p_0	0.1	Hill constant for mRNA transcription pf Hes1
α	0.5	partition term between NICD and Hes1 protein effect on Hes1 mRNA
b_H	0	leackiness for Hes1 promoter
b_p	24300	maximal transcription rate for the reporter gene
K_{RS}	300000	Hill constant for Notch action on Hes1
b_N	40	constant production for Notch
b_D	40	maximal transcription rate for Delta ligand
d_D	0.1	degradation of Delta ligand
K_p	0.01	constant Hill for Delta
K_f	0.135	protein folding rate

Table 8.2: Table of parameters for the three proposed models

The LIMIR model can be obtained by adding the term $-\frac{N_i D_i}{K_c}$ to Eqs. 8.37 and 8.38. The function f is periodic and depends on the term $N_i D_j$.

8.5.1 The decoupled system $\epsilon = 0$

Let us consider for it only the single cell system (i.e. no coupling, which implies $\epsilon = 0$). In this case, Eqs. 8.37 and 8.38 become:

$$\frac{dN}{dt} = b_N - N \tag{8.39}$$

$$\frac{dD}{dt} = b_D \cos(\omega t) - D. \tag{8.40}$$

The analytical solution is easy to obtain:

$$N(t) = b_N - (b_N - N_0)e^{-t} \tag{8.41}$$

$$D(t) = \left(D_0 - \frac{b_D}{1+\omega^2}\right) e^{-t} + \frac{b_D(\cos(\omega t) + \omega \sin(\omega t))}{1+\omega^2}. \tag{8.42}$$

So the term $D(t)$ is always an oscillating function with a transient phase given by the exponential term.

When $t \to +\infty$, $D(t)$ can be approximated with the term $\frac{b_D}{1+\omega^2}[\cos(\omega t) + \omega \sin(\omega t)]$. This term can be rewritten as $\alpha \cos(\omega t - \beta)$. This implies

$$\alpha \cos\beta = \frac{b_D}{1+\omega^2} \alpha \sin\beta = \frac{b_D \omega}{1+\omega^2}$$

and then

$$\tan\beta = \frac{1}{\omega} \text{ and } \alpha = \frac{b_D}{\sqrt{1+\omega^2}}.$$

8.5.2 Single cell with an external constant input ($\epsilon = const.$)

Now we model the case of a single cell with an external constant input ϵ:

$$\frac{dN}{dt} = b_N - N - \epsilon \tag{8.43}$$

$$\frac{dD}{dt} = b_D \cos(\omega t + \frac{\epsilon}{K+\epsilon}) - D. \tag{8.44}$$

As before, the equations are decoupled and the analytical solution is easy to obtain:

$$N(t) = b_N - \epsilon + (N_0 - b_N + \epsilon)e^{-t} \tag{8.45}$$

$$\begin{aligned} D(t) &= \left(D_0 - \frac{b_D}{1+\omega^2} - b_D\frac{\epsilon}{K+\epsilon}\right)e^{-t} \\ &+ \frac{b_D(\cos(\omega t) + \omega\sin(\omega t))}{1+\omega^2} + b_D\frac{\epsilon}{K+\epsilon}. \end{aligned} \tag{8.46}$$

The general behaviour of the system is the same as before: in particular, the model structure can be reduced to the previous case as shown in the following:

$$\overline{N} = N + \epsilon \tag{8.47}$$

$$\overline{D} = D - \frac{b_D\epsilon}{K+\epsilon}; \tag{8.48}$$

so the equations become as in the previous case:

$$\frac{d\overline{N}}{dt} = b_N - \overline{N} \tag{8.49}$$

$$\frac{d\overline{D}}{dt} = b_D \cos(\omega t) - \overline{D}. \tag{8.50}$$

8.5.3 Single cell with an external time-varying input ($\epsilon = \epsilon(t)$.)

In the case of a single cell with an external input $\epsilon = \epsilon(t)$, the equations to be considered are:

$$\frac{dN}{dt} = b_N - N - \epsilon(t) \tag{8.51}$$

$$\frac{dD}{dt} = b_D \cos(\omega t + \frac{\epsilon(t)}{K + \epsilon(t)}) - D. \tag{8.52}$$

The analytical solution is:

$$N(t) = b_N - e^{-t} \int e^t \epsilon(t) dt + (N_0 - b_N + \int \epsilon(t) dt) e^{-t} \tag{8.53}$$

$$\begin{aligned} D(t) &= Ce^{-t} + b_D e^{-t} \int \frac{e^t \cos(\omega t)\epsilon(t)}{K + \epsilon(t)} dt + b_D e^{-t} \int \frac{e^t \epsilon(t)}{K + \epsilon(t)} dt \\ &+ K b_D e^{-t} \int \frac{e^t \cos(\omega t)}{K + \epsilon(t)} dt, \end{aligned} \tag{8.54}$$

where C is a constant obtained by setting $D(0) = D_0$. It is now possible to study the response of the oscillator to a constant, periodic or oscillating input. Due to the complexity of the analytical form and in order to obtain explicit

analytical results, I considered as a starting point the following model, in which the influence of the external stimulus has been simplified:

$$\frac{dN}{dt} = b_N - N - \epsilon(t) \tag{8.55}$$

$$\frac{dD}{dt} = b_D \cos(\omega t + \epsilon(t)) - D. \tag{8.56}$$

The solution for this model now becomes:

$$N(t) = b_N - e^{-t} \int e^t \epsilon(t) dt + (N_0 - b_N + \int \epsilon(t) dt) e^{-t} \tag{8.57}$$

$$D(t) = Ce^{-t} + b_D e^{-t} \int e^t \epsilon(t) dt + \frac{b_D}{1+\omega^2} [\cos(\omega t) + \omega \sin(\omega t)], \tag{8.58}$$

where

$$C = D_0 - \frac{b_D}{1+\omega^2} - b_D \left[\int e^t \epsilon(t) dt \right]_{t=0}.$$

8.5.4 The case $\epsilon(t) = \sin(\omega t)$

When $\epsilon(t) = \sin(\omega t)$, the equation for $N(t)$ its the same as in 8.37, while $D(t)$ becomes:

$$\frac{dD}{dt} = b_D(\cos(\omega t) + \sin(\omega t)) \tag{8.59}$$

so the analytical solution becomes:

$$D(t) = Ce^{-t} + \frac{b_D}{1+\omega^2} [(\sin \omega t + \cos \omega t) + \omega(\sin \omega t - \cos \omega t)]. \tag{8.60}$$

Otherwise, the equation can be rewritten as

$$\frac{dD}{dt} = b_D(\cos(\omega t) + \sin(\omega t)) \tag{8.61}$$

$$= b_D(\alpha \cos(\omega t - \beta)) - D \tag{8.62}$$

8.5.5 The case $\epsilon(t) = A\cos(\omega t) + B\sin(\omega t)$

Let us assume that our oscillator is perturbed by a generic signal $\epsilon(t) = A\cos(\omega t) + B\sin(\omega t)$, hence the equation D becomes:

$$\frac{dD}{dt} = b_D(\cos(\omega t) + (A\cos(\omega t) + B\sin(\omega t))) - D \tag{8.63}$$

$$= b_D((1+A)\cos(\omega t) + B\sin(\omega t)) - D \tag{8.64}$$

This form of the external input can be rewritten as

$$\epsilon(t) = \alpha \cos(\omega t - \beta) \tag{8.65}$$

where, $\alpha = \sqrt{((1+A)^2 + B^2)}$ and $\tan(\beta) = \frac{B}{1+A}$.

The equation becomes:

$$\frac{dD}{dt} = b_D(\sqrt{((1+A)^2 + B^2)}\cos\left(\omega t - \arctan\frac{B}{1+A}\right)) - D. \tag{8.66}$$

So, solving this kind of problem means solve the problem with the same oscillator with a different amplitude in oscillation and phase shift of arctan $\frac{B}{1+A}$. The analytical solution is

$$D(t) = Ce^{-t} + \frac{b_D}{1+\omega^2}\sqrt{(\omega + A\omega + B)^2 + (1 + A - B\omega)^2}\cos(\omega t - \arctan\frac{\omega + A\omega + B}{1 + A - B\omega}). \tag{8.67}$$

In conclusion, when perturbed by an harmonic oscillation, the resulting oscillation is always shifted and modified in amplitude.

8.6 Analytical study on the LIR and LIMIR models for the $\mathcal{N}$ cell scenario

After considering the case of the single cell, I analysed the same model this time extended to the $\mathcal{N}$ cells using methods from the perturbation theory.

Let us assume that our coupling term is $\epsilon << 1$, we can study the asymptotic behaviour of the system (Eqs. 8.37 and 8.38) starting from the synchroneous solution by adding a "small" perturbation term, in order to obtain the approximate solution.

The synchroneous solution is:

$$\frac{dN}{dt} = b_N - N - \epsilon ND \tag{8.68}$$

$$\frac{dD}{dt} = b_D(\cos(\omega t) + \epsilon ND) - D - \epsilon ND \tag{8.69}$$

Let us assume that N, D can be rewritten as:

$$N = N_0 + \epsilon N_\epsilon + O(\epsilon^2) \tag{8.70}$$

$$D = D_0 + \epsilon D_\epsilon + O(\epsilon^2), \tag{8.71}$$

so that:

$$\dot{N} + \epsilon\dot{N}_\epsilon = b_N - N_0 - \epsilon N_\epsilon - \epsilon N_0 D_0 \tag{8.72}$$

$$\dot{D} + \epsilon\dot{D}_\epsilon = b_D(\cos(\omega t) + \epsilon N_0 D_0) - D_0 - \epsilon D_\epsilon - \epsilon N_0 D_0 \tag{8.73}$$

Solving this system, the analytical solution becomes:

$$\begin{aligned} N(t) &= b_N - C_1 e^{-t} + \epsilon C_3 e^{-t} - \epsilon e^{-t}[\frac{t}{\sqrt{1+\omega^2}}(C_2 b_N \sqrt{1+\omega^2} \\ &- C_1 b_D \cos(\omega t - \arctan\frac{1}{\omega})) + C_1 C_2 e^{-t} \\ &+ \frac{b_D b_N e^t}{\sqrt{1+\omega^2}} \cos(\arctan\frac{1}{\omega} - \omega t)] \end{aligned} \tag{8.74}$$

$$\begin{aligned} D(t) &= C_2 e^{-t} + \frac{b_D}{\sqrt{1+\omega^2}} \cos(\omega t - \arctan\frac{1}{\omega}) \\ &+ C_4 e^{-t} + \frac{(b_D - 1)e^{-2t}}{\sqrt{1+\omega^2}}[\frac{C_1 C_2}{\sqrt{1+\omega^2}} \\ &+ b_D b_N e^{-2t} \cos(\omega t - \arctan\frac{1}{\omega}) + \frac{C_2 b_N t e^t}{\sqrt{1+\omega^2}} \\ &- C_1 b_D t e^t \cos(\omega t - \arctan\frac{1}{\omega})]. \end{aligned} \tag{8.75}$$

Thus a synchroneous oscillatory behaviour is possible also with a weak coupling.

8.6.1 A numerical analysis of the LIR and LIMIR models for the $\mathcal{N}$ cells scenario

Let us consider the equations for the LI model for the "i-cell" derived in Section 8.4 which I report again for clarity:

$$\frac{dN_i}{dt} = \tau(b_N - N_i - \sum_{j=]i[} D_j N_i) \tag{8.76}$$

$$\frac{dD_i}{dt} = \tau(b_D \frac{1}{1+{p_{2i}}^2} - \sum_{j=]i[} N_j D_i - D_i) \tag{8.77}$$

$$\frac{dR_i}{dt} = \tau(b_R \left(\frac{(N_i \sum_{j=]i[} D_j)^n}{K_{RS} + (N_i \sum_{j=]i[} D_j)^n} \right) \left(\alpha + (1-\alpha) \frac{k{p_0}^2}{{p_0}^2 + {p_{2i}}^2} \right) - R_i) \tag{8.78}$$

$$\frac{dp_{1i}}{dt} = \tau(b_p R_i - (1+K_f) p_{1i}) \tag{8.79}$$

$$\frac{dp_{2i}}{dt} = \tau \left(K_f p_{1i} - \frac{p_{2i}}{K_p + p_{2i}} \right). \tag{8.80}$$

and similarly for the equations for the LIMI model:

$$\frac{dN_i}{dt} = \tau(b_N - N_i - \sum_{j=]i[} D_j N_i - N_i D_i) \tag{8.81}$$

$$\frac{dD_i}{dt} = \tau(b_D \frac{1}{1+{p_{2i}}^2} - \sum_{j=]i[} N_j D_i - D_i - N_i D_i) \tag{8.82}$$

$$\frac{dR_i}{dt} = \tau(b_R \left(\frac{(N_i \sum_{j=]i[} D_j)^n}{K_{RS} + (N_i \sum_{j=]i[} D_j)^n} \right) \left(\alpha + (1-\alpha) \frac{k{p_0}^2}{{p_0}^2 + {p_{2i}}^2} \right) - R_i) \tag{8.83}$$

$$\frac{dp_{1i}}{dt} = \tau(b_p R_i - (1+K_f) p_{1i}) \tag{8.84}$$

$$\frac{dp_{2i}}{dt} = \tau \left(K_f p_{1i} - \frac{p_{2i}}{K_p + p_{2i}} \right). \tag{8.85}$$

$$\tag{8.86}$$

with symbol $]i[$, I indicate the set of the indexes defining the six cells neighbouring of the cell i in the hexagonal lattice in Figure 7.10. All the components of the system have the same meaning of the system for one cell with the extra of the contribution of the other cells (the sum), thus obtaining a system of $5 \times \mathcal{N}$ equations.

In order to numerically study the behaviour of a network composed by $\mathcal{N}$ cells, I considered methods similar to those developed for patterning in Chapter 7: the collective behaviour of the population of coupled oscillators can be simulated using the connectivity matrix and periodic boundary conditions to simulate a tissue of cells each one interacting with six neighbouring cells. Differently from patterning, now over time the tissue does not approach a

static configuration, but a common solution, in the case of synchronization, or present other dynamical behaviours depending on the choice of parameters. In order to obtain a plot of the region of parameters in which synchronization can occur, I performed an analysis on the Master Stability Function as in [108], [1].

The Master Stability Function (MSF) is an useful instrument able to detect the convergence of a network of oscillators to a stable limit cycle. The attractive state is the synchroneus state where all the cells are equal, thus presenting the same states over time. Starting from

$$\frac{d\underline{x}}{dt} = F(\underline{x})$$

as the equation of the network where each $\underline{x}$ represents the dynamical variable describing the single cell, the synchroneous state is

$$\underline{s} = \underline{x}_1 = \underline{x}_2 = ... = \underline{x}_{\mathcal{N}};$$

thus, it is possible to compute the largest Lyapunov exponent from the Jacobian obtaining the Maximum Stability Function of the system.

Only when MSF is negative, the oscillators will synchronize. In this way, I obtained a parameters' region where synchronization can occur as shown in Figures 8.19 and 8.20. Figure 8.20 shows the absolute values of the MSF while, in Figure 8.19, I rescaled the values in order to more easily compare the results for LIR and LIMIR models. In both models, when $n = 1$ the region of synchronization is larger with respect to the case $n = 3$. Figures 8.21 and 8.22

illustrate two representative simulations of the LIR model for an hexagonal lattice of 9 cells. Figure 8.21 shows the numerical solution for the variable N_i for all the 9 cells, when b_N and b_D are chosen in the synchronisation region. It can be observed that each oscillator start from a different initial condition but reach the same oscillatory solution (synchronization case), as expected. Figure 8.22 shows the numerical solution when b_D and b_N are chosen outside of the synchronization region. In this case, cells are not able to reach a common behaviour (no-synchronization case), as expected.

In order to better appreciate the different behaviours in the case of a non-synchronized behaviour, I show representative simulations for the LIR model for a lattice of 12×12 cells with paramters b_D and b_N have value outside the synchronization parameter region, as reported in Figs. 8.23, 8.24, 8.24, 8.25 and 8.26.

In the case of the LIMIR model, similar behaviours can be obtained,

as reported in Figures 8.28, 8.30, 8.30 for $n = 1$ and 8.31, 8.32, 8.33 for $n = 3$.

In conclusion, we can compare the uniform/patterning zone for LI/LIMI models 7.11, with the synchronization/unsynchronization region for LIR/LIMIR models 8.19, thus observing that parameters in which synchronization of cells can occur are not the same in which patterning can occur.

Comparing results from the Maximum Lyapnov Exponent in the case of a non-oscillatory gene (Figure 7.11) and the Maximum Stability Function in

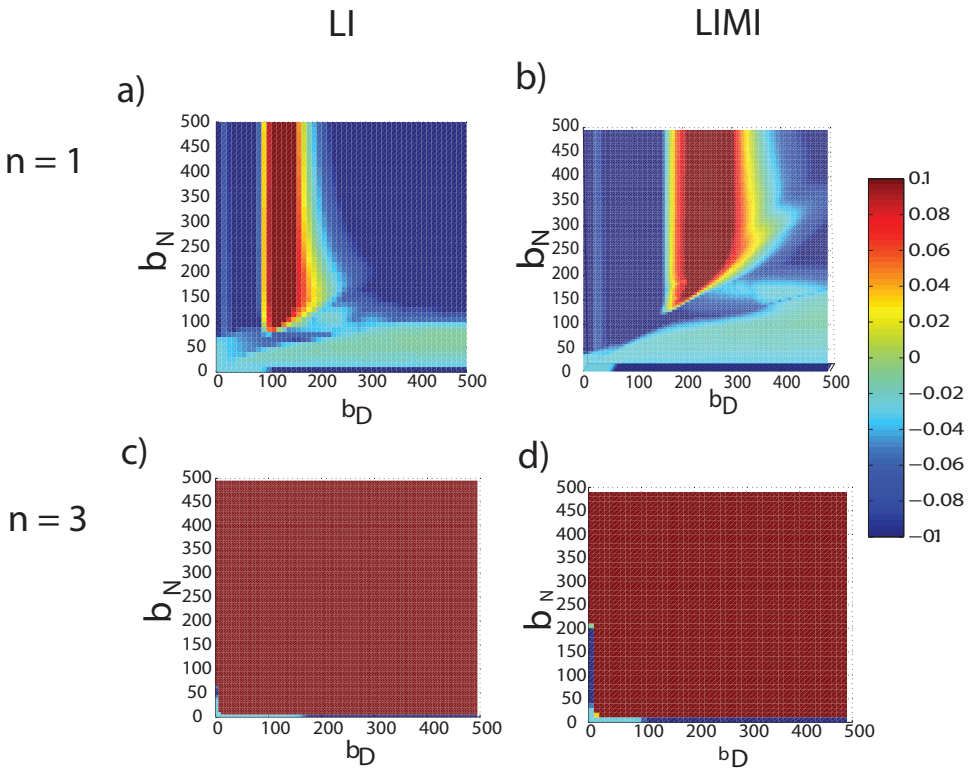


Figure 8.19: **Master Stability Function.** MSF rescaled to the interval $[-0.1, 0.1]$ for (a) LIR with $n = 1$ (b) LIMIR with $n = 1$ (c) LIR with $n = 3$ (d) LIMIR $n = 3$.

the case of an oscillatory gene (Figure 8.19), it is possible to see that the synchronization region in the MSF does not overlap with the patterning region in the MLE plot. This implies that when the expression of the oscillatory genes is synchronized, the system cannot approach a pattern when the negative feedback is silenced: synchronization of oscillators implies an uniform steady state configuration when the feedback stops.

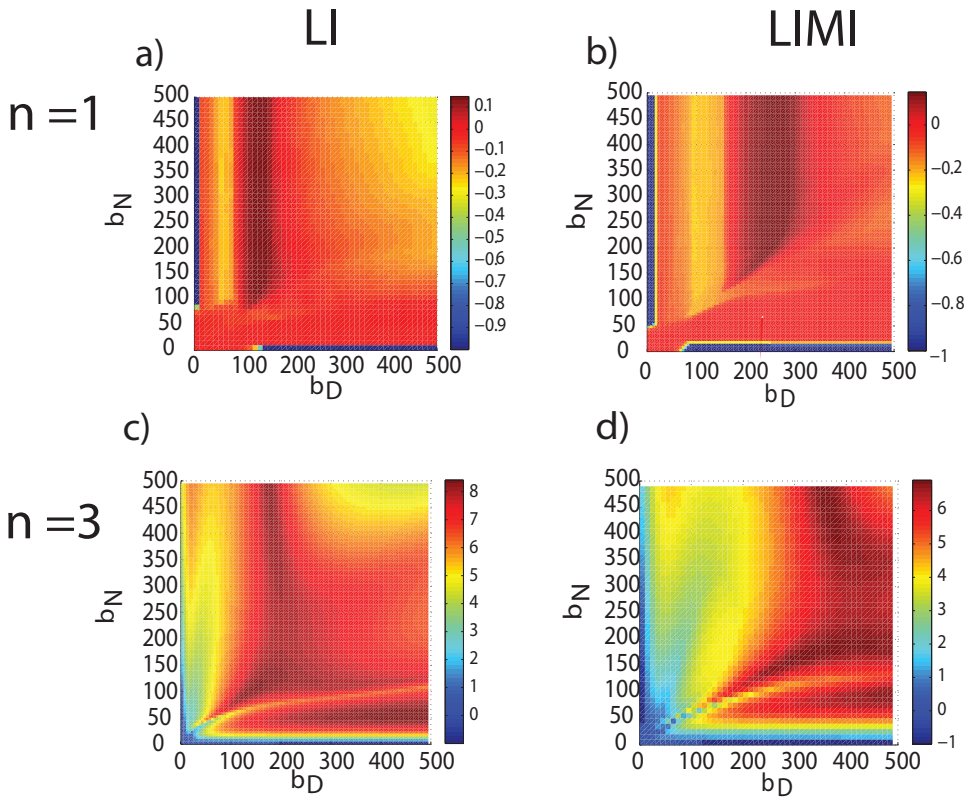


Figure 8.20: **Master Stability Function.** MSF absolute values for (a) LIR with $n = 1$ (b) LIMIR with $n = 1$ (c) LIR with $n = 3$ (d) LIMIR $n = 3$.

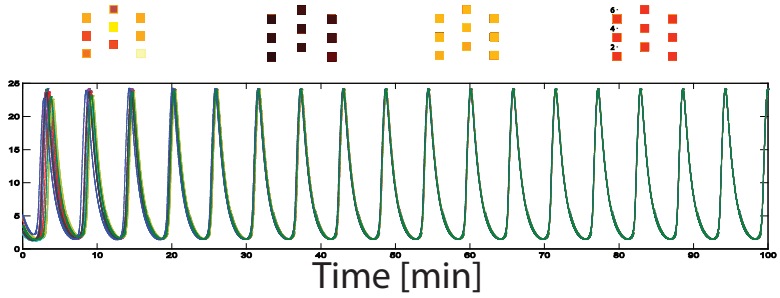


Figure 8.21: **Synchronized cells.** Simulated network of 3×3 cells starting from different initial conditions and synchronizing. Above inlets are illustrative for the network dynamics: each cell is represented by a square and can communicate with 6 neighbouring cells, each one reaching different levels of expression first and then the same.

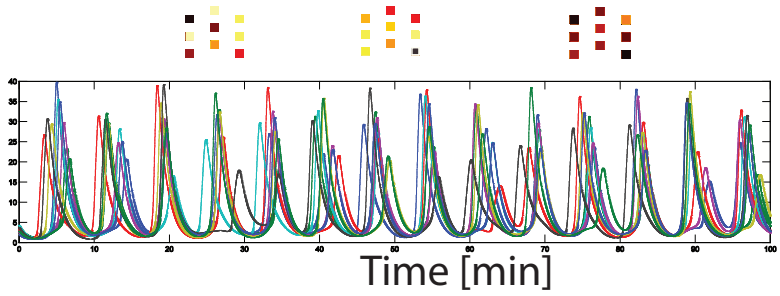


Figure 8.22: **Unsynchronized cells.** Simulated network of 3×3 cells starting from different initial conditions and continuing to express at random different levels. Above inlets are illustrative for the network dynamics: each cell is represented by a square and can communicate with 6 neighbouring cells, each one reaching different levels of expression.

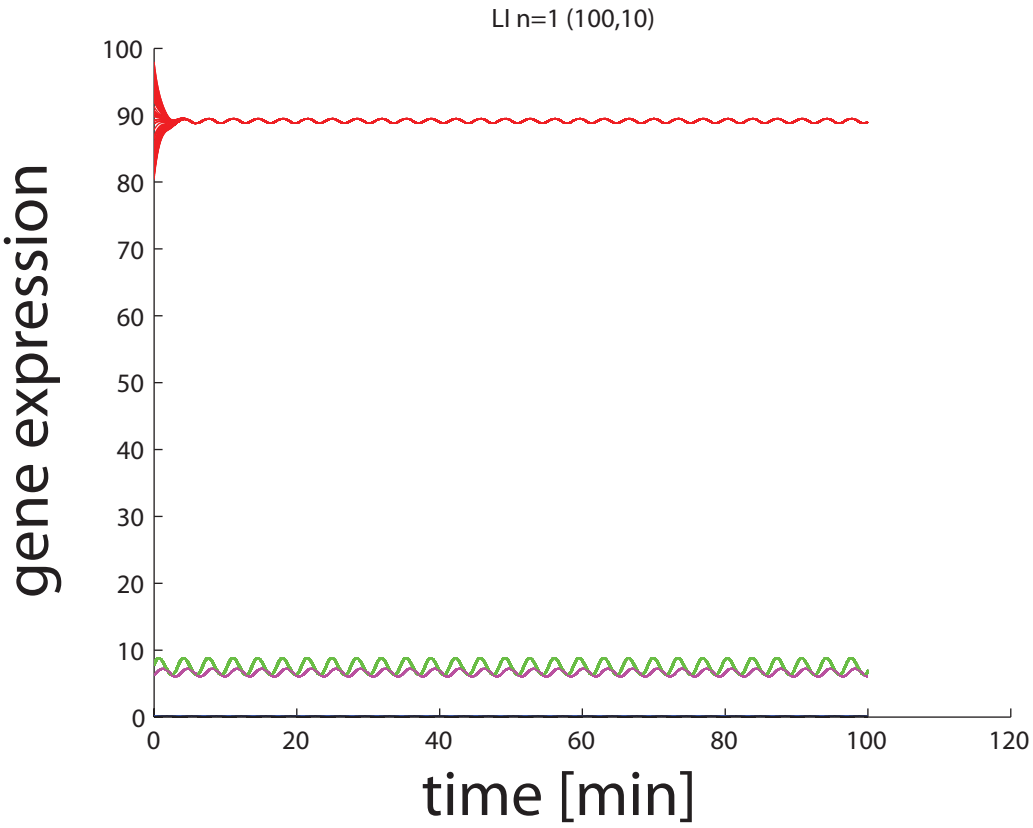


Figure 8.23: **Simulation for LIR.** Expression in time for a LIR system in which $n = 1$ and b_D and b_N are 100 and 10. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

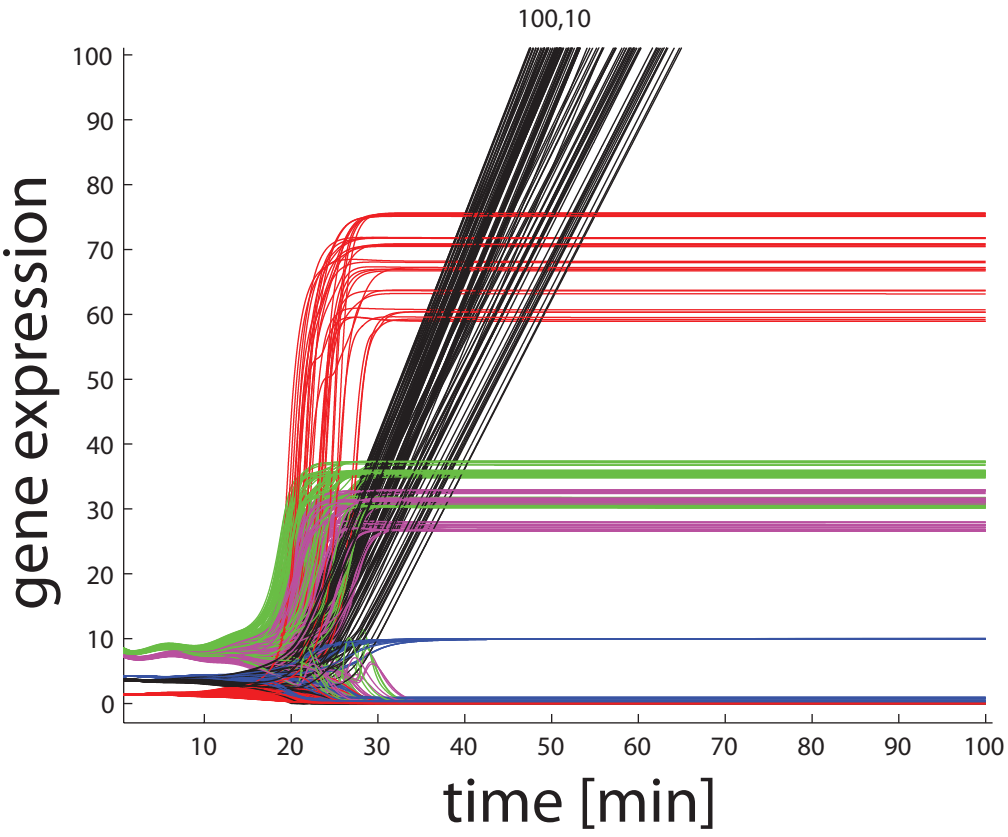


Figure 8.24: **Simulation for LIR.** Expression in time for a LIR system in which $n = 3$ and b_D and b_N are 100 and 10. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

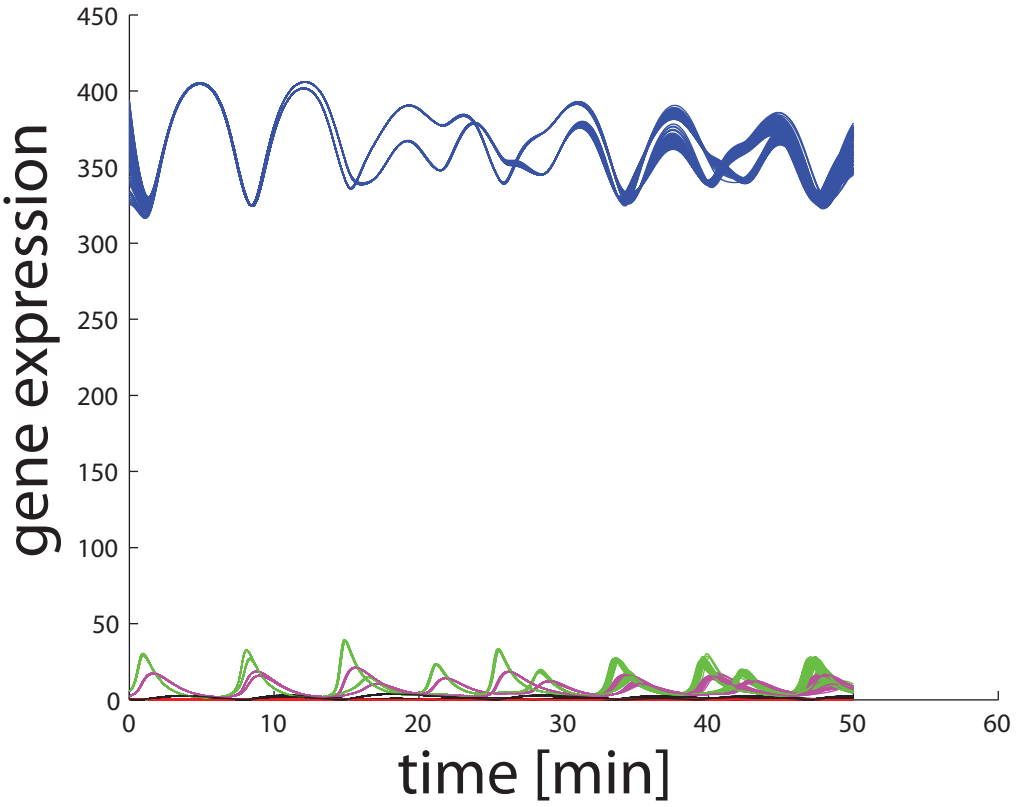


Figure 8.25: **Simulation for LIR.** Expression in time for a LIR system in which $n = 1$ and b_D and b_N are 130 and 430, $MLE = 0.14$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

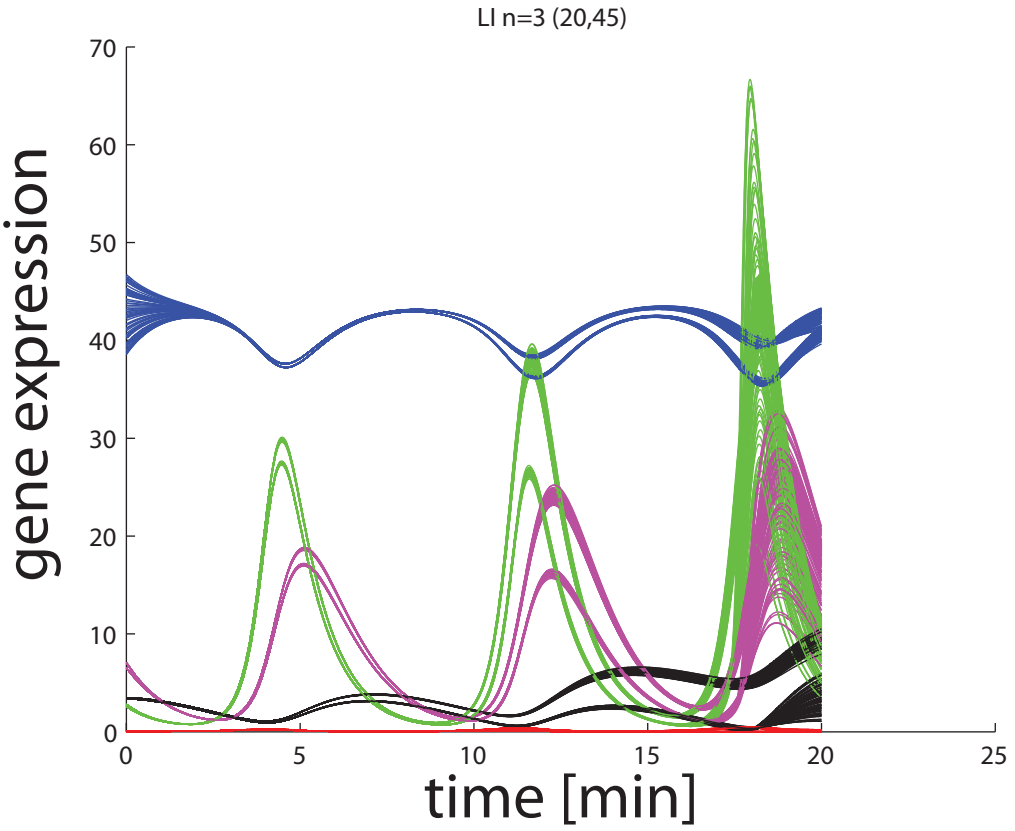


Figure 8.26: **Simulation for LIR.** Expression in time for a LIR system in which $n = 3$ and b_D and b_N are 20 and 45, $MLE = 2.9$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

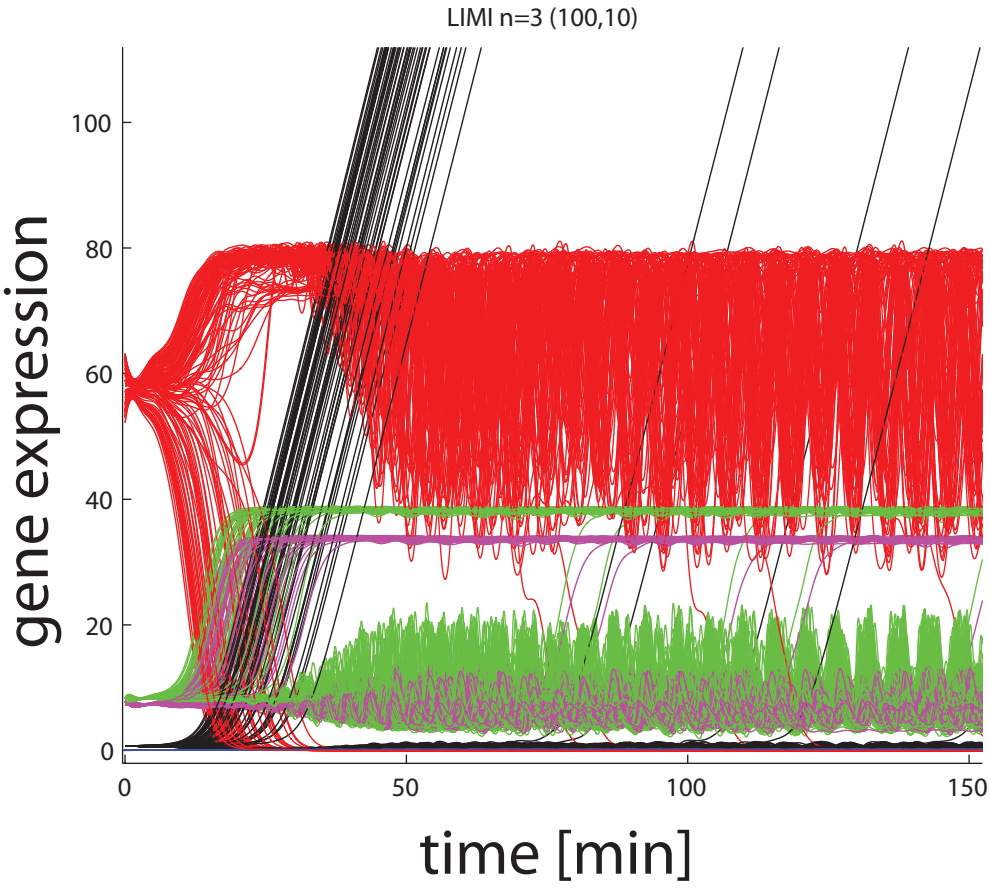


Figure 8.27: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 1$ and b_D and b_N are 100 and 10. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

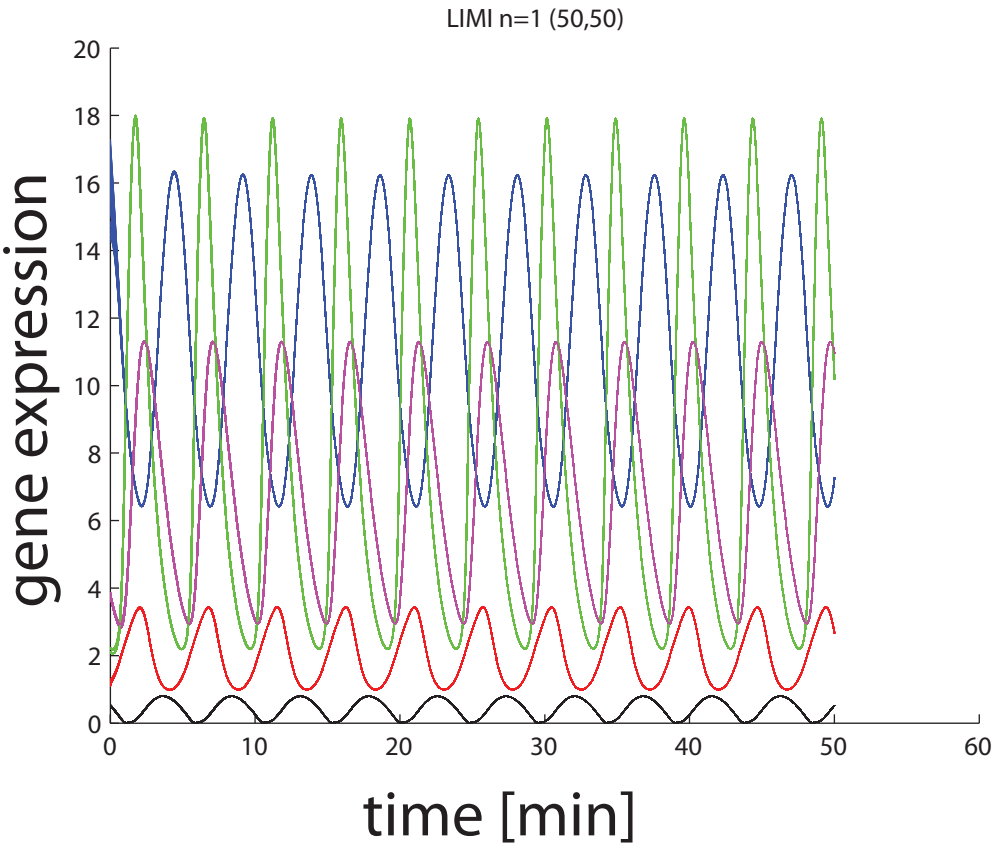


Figure 8.28: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 1$ and b_D and b_N are 50 and 50, $MLE = -0.048$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

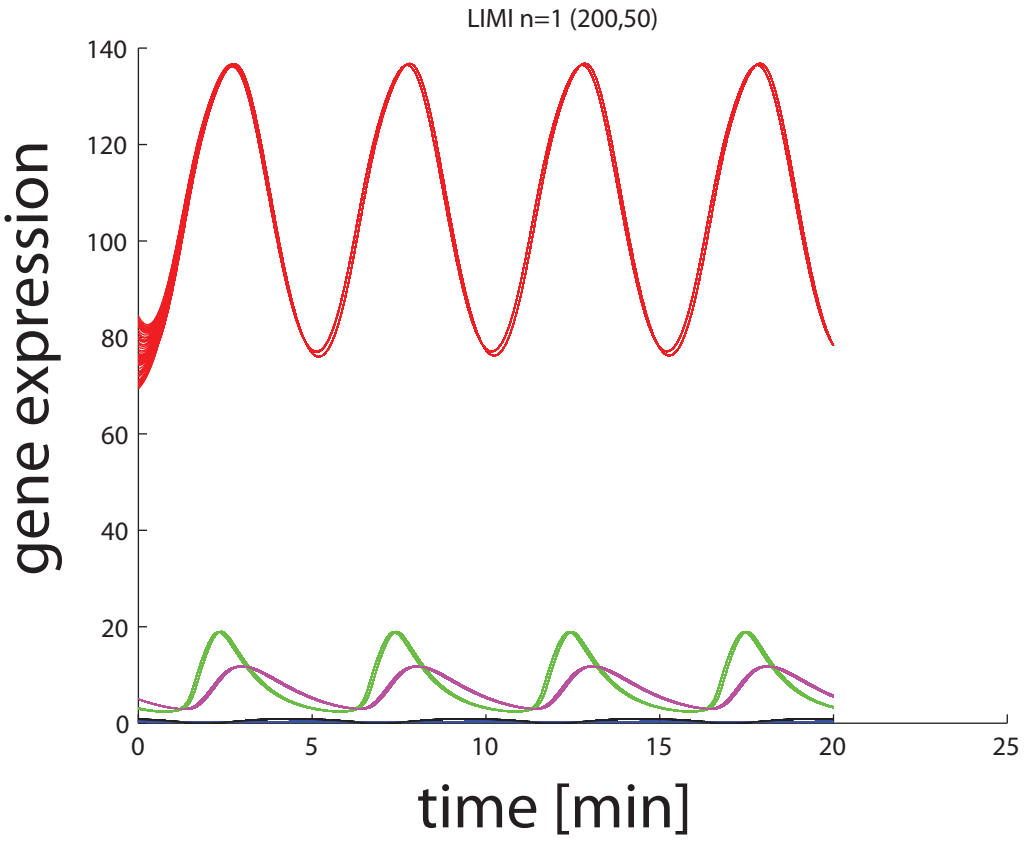


Figure 8.29: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 1$ and b_D and b_N are 200 and 50, $MLE = -0.0234$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

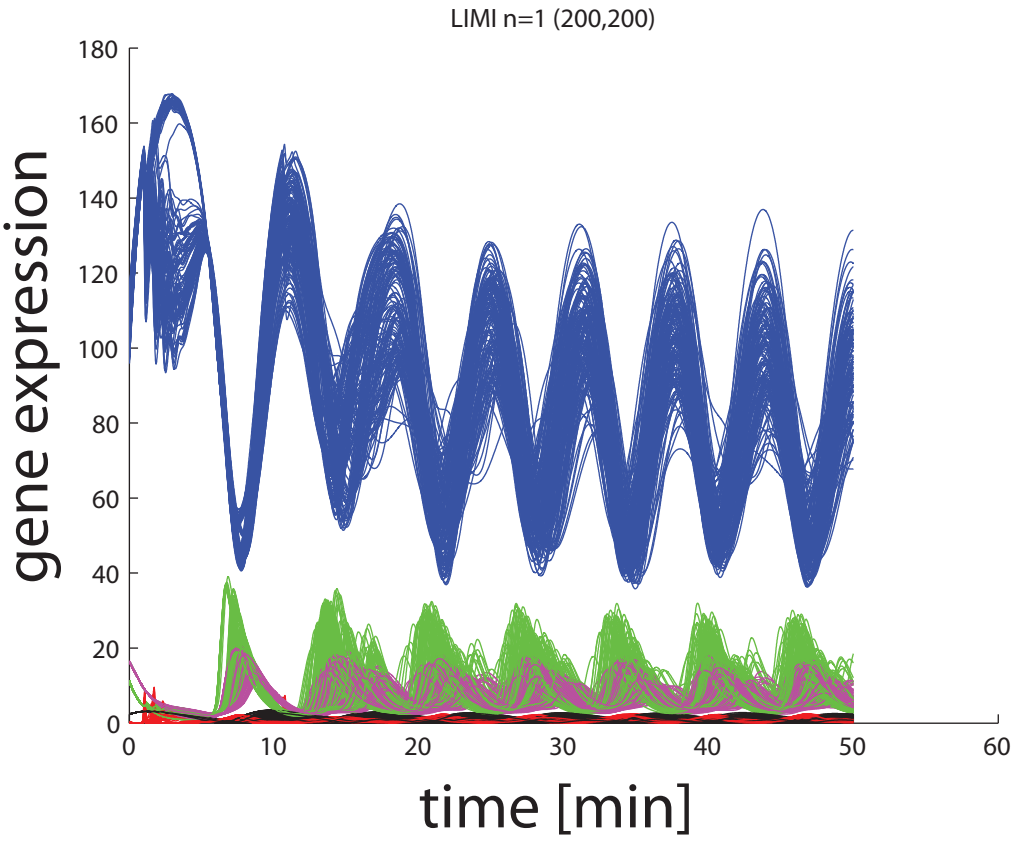


Figure 8.30: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 1$ and b_D and b_N are 200 and 200, $MLE = 0.1439$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

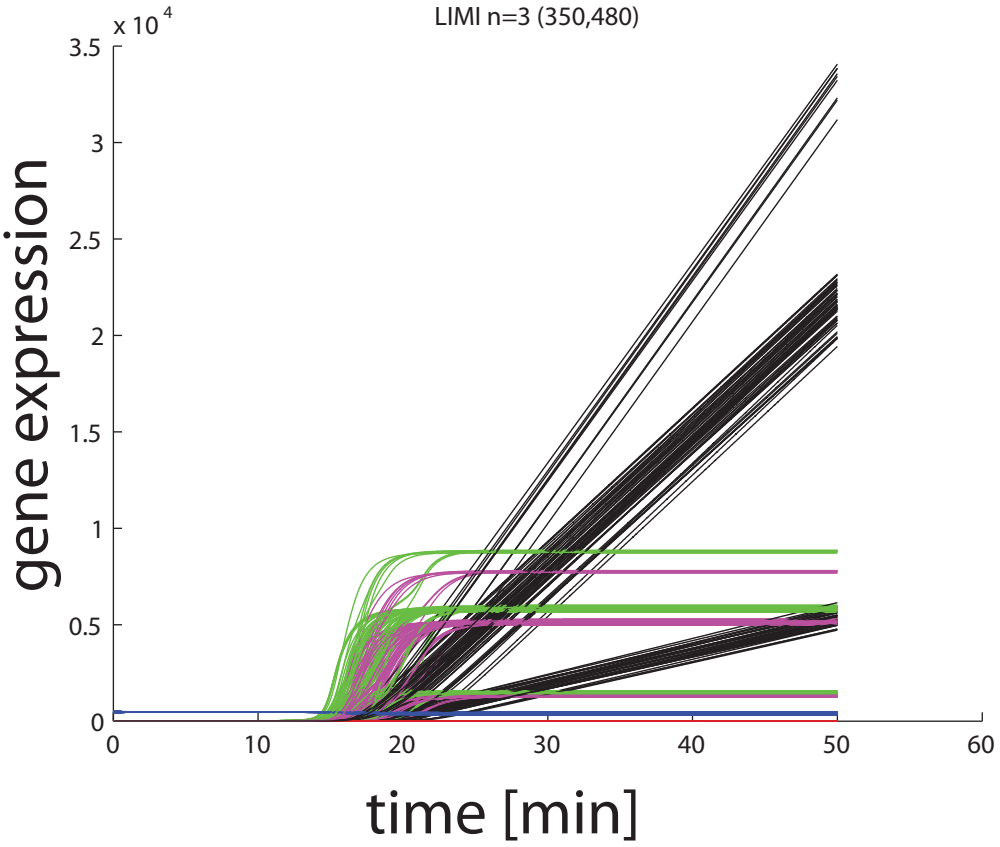


Figure 8.31: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 3$ and b_D and b_N are 350 and 480, $MLE = 6.89$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

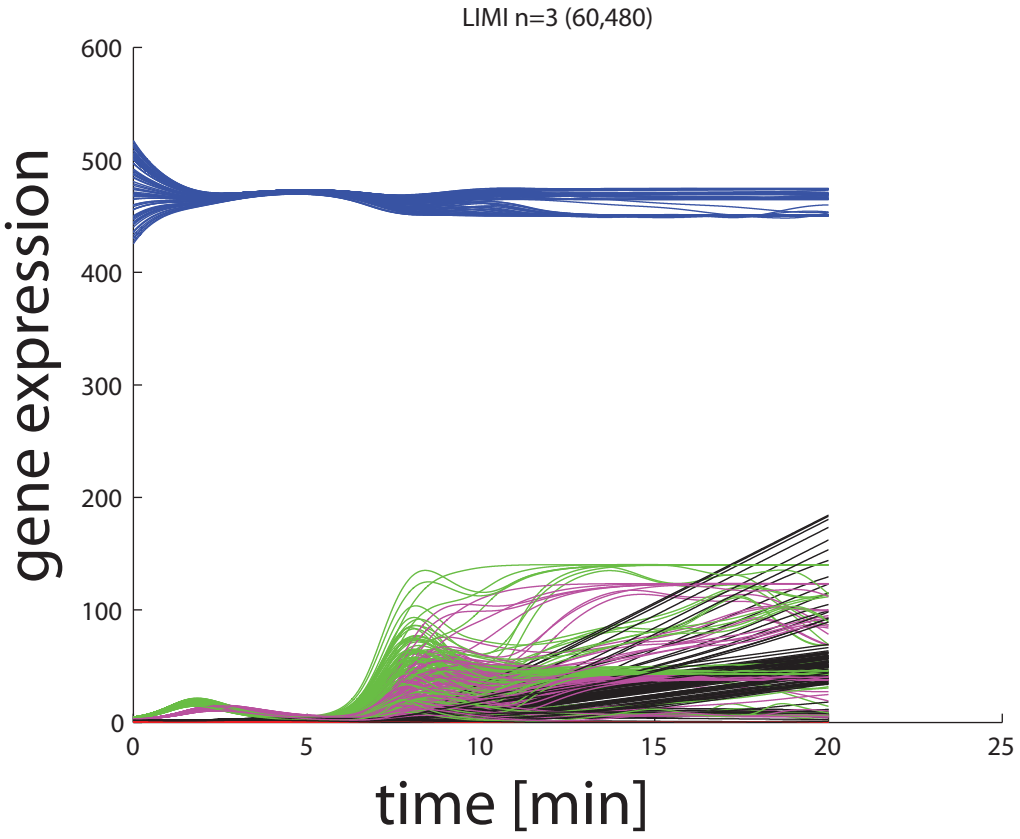


Figure 8.32: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 3$ and b_D and b_N are 60 and 480, $MLE = 4.25$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

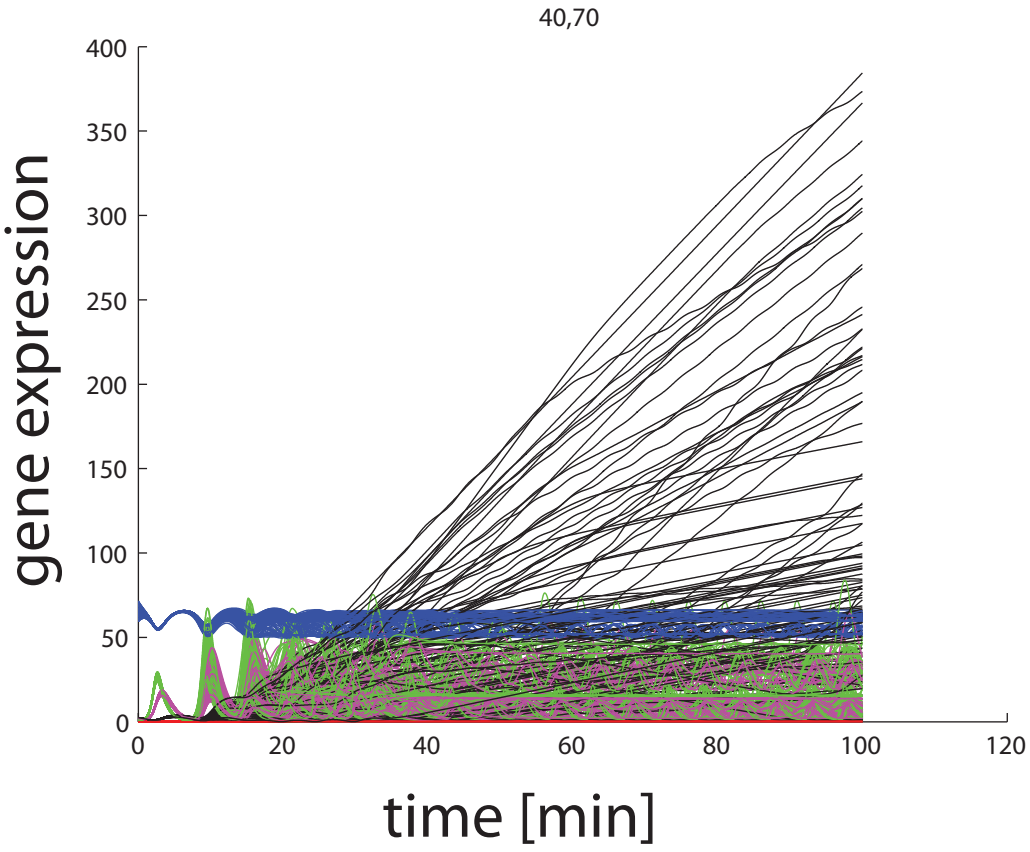


Figure 8.33: **Simulation for LIMIR.** Expression in time for a LIMIR system in which $n = 3$ and b_D and b_N are 40 and 70, $MLE = 1.16$. (Blue = Notch, Red = Dll, Green = mRNA reporter, Magenta = unfolded protein reporter, Black = folded protein reporter)

Chapter 9

Modeling *Synchronization*: the somite segmentation clock

Here, the model of synchronization as presented in Chapter 8 is adapted to study a more specific process: somitogenesis. Specifically, the Notch-Delta-Hes model, derived in Chapter 8, is used to study somite formation during embryogenesis by simulating the transcriptional wave propagation across the pre-somitic mesoderm observed during vertebrate embryogenesis.

9.1 Somitogenesis timing and boundary formation

As explained in Chapter 6, somitogenesis is the process in which somites form during the embryo elongation. The periodicity of somites is evident in all vertebrate species and their features have been linked to the activity of the autonomous oscillations of the Hes gene family in mammalian cells. During the somites development, the Notch pathway couples the different cells, thus allowing the synchronization of oscillators. Vertebrae size and number can be correlated to the period of this oscillation. It has been observed experimentally that, during this phase of embryo development, PSM cells are able to propagate a periodic, anteriorly traveling wave of cyclic gene expression [12]. The signal of the segmentation clock is so translated into a spatial, periodic pattern by a complex signalling gradient system within the presomitic mesoderm (PSM).

The *Clock and Wavefront model*, as explained in Chapter 6, is the actual accepted model of somite segmentation. However, recent studies show that somitogenesis is an oscillator-independent process (see [106], [43]): Notch is considered to be necessary for the synchronization of the genetic oscillators and for the positioning of new somites, but somites are seen to form also in absence of Hes cyclic expression. In this chapter, I provide an alternative hyphotesis of how a transcriptional wave can travel across the tissue. Specifi-

cally, I considered the hyphotesis that few cells are responsible for starting the oscillatory behaviour and this stops at the correct position without requiring a gradient of some molecule as for the *Clock and Wavefront model*.

9.2 A new model for the somite segmentation clock

In the *Clock and Wavefront* model, the propagation of a wave along the embryo is considered a process necessary for somite formation: to obtain this traveling wave, the cells are assumed to be autonomous oscillators. Here, I will show by simulation that a wave can travel along a stripe when all the PSM cells are autonomous oscillators, according to the classical model: a short pulse given to the first "group" of oscillating cells (first column of the stripe) is able to propagate to the other adiacent cells in an ordered way (Section 9.3). These propagation and shifting of phase is transient and then oscillators came back in phase to the synchroneous solution.

However, according to [43], the formation of somites can be considered as a process independent of the oscillatory behaviour: indeed the authours demonstrated that isolating a portion of PSM of a quail embryo in a peripherical zone of it and treating the cells with an antagonist of the bone morphogenetic protein (BMP), somites can form in grapes without showing an oscillatory behaviour of the Hes. In Figure 9.1, the experimental setup is

shown and the grape of resulting somites is indicated.

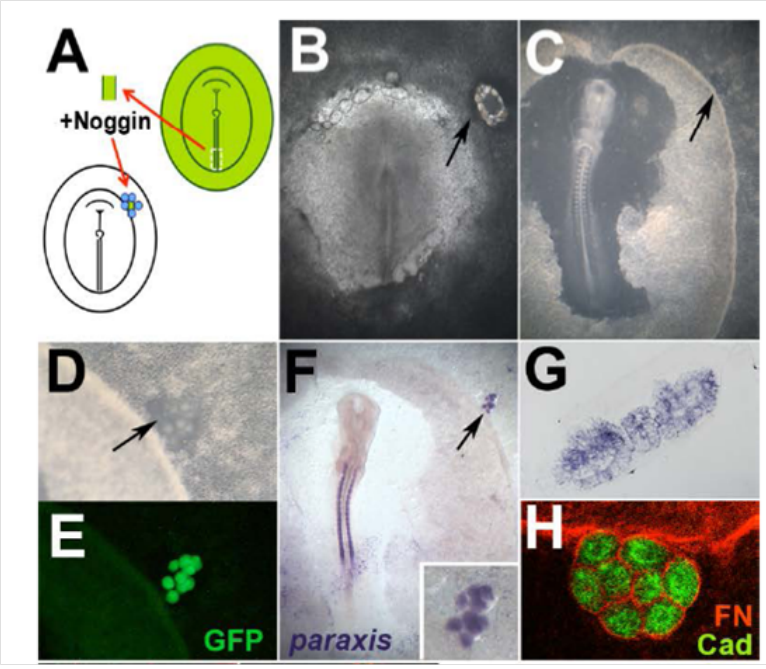


Figure 9.1: **BMP inhibition generates normal somites. From ([43]).** Experimental design. The PS of a donor quail or GFP-transgenic embryo is excised; exposed to Noggin (BMP antagonist); and grafted, surrounded by Noggin-beads, to the periphery of a host chick embryo [(A and B), arrows]. After overnight incubation, a group of somite-like structures arranged as a bunch of grapes appears [(C and D), arrows]. These structures fluoresce if the donor is a GFP-transgenic embryo (E). The ectopic structures are real somites: They express paraxis (F and G) and N-cadherin [green in (H)] and are surrounded by a Fibronectin matrix [red in (H)].

The formed somites express all the somite characteristics except for the characteristic rostro-caudal polarization. Hence, the data suggest that somites can be thought as self-organizing structures more related to a phenomenon of "patterning" rather to synchronization regime.

Therefore, it is possible to assume that the traveling wave is responsible for the correct timing and spacing of somites and of their rostro-caudal polarity, but not of their formation. Thus, I simulated a set of scenarios in which cells do not exhibit autonomous oscillations. Biologically, this hypothesis implies that cells in the presomitic mesoderm are not able to express cyclically the Hes gene autonomously but require a group of pacemaker cells to start and sustain the oscillations. Tha advantage of this modelling approach is that, in the case the cell can present dampened oscillations when externally entrained, no gradient is required to stop the wave propagation, differently from the *Clock and Wavefront* model, as shown in what follows.

9.3 Oscillating PSM cells scenario

First, I performed a series of numerical simulations to show that, in the regime of synchronization, it is possible to obtain a transient wave propagation along a stripe of cells by transiently perturbing one column of cells, at one of the edges of the stripe. A traveling wave of this kind has been observed during the somitogenesis process. Indeed, during the elongation of the embryo a wave of gene expression starting from the tail and directed to the anterior part is

observed. In order to simulate the stripe of cells, I imposed fixed boundary conditions so that the only cells receiveing the stimulus are those of the first column. The stripe considered has a length of 20 cells and an height of 3 as shown in Figure 9.2. The input given to the first column of three cells is a pulse applied to all the five state variables of the LIR (or LIMIR) model. An example of propagation over space and time along the stripe is shown in Figure 9.2 and Figure 9.3, where it is possible to estimate the effects of the transient shifting in phase. Of course there are several pulses that can give rise to this type of propagation, depending on the starting point of the pulse and its duration.

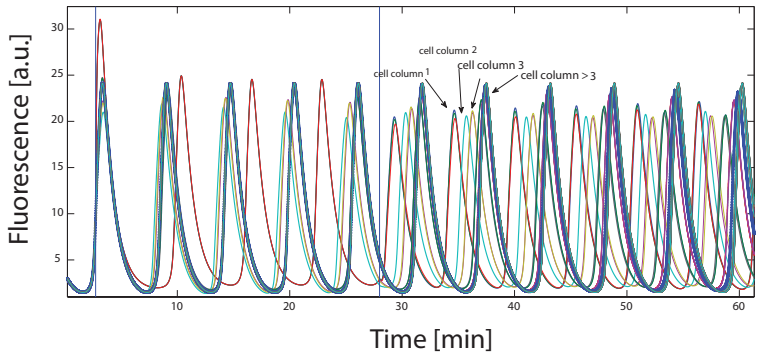


Figure 9.2: **Propagation of a signal along a stripe, over time.** A pulse (thin blue line) is given to the synchronized system of 20×3 cells: after removing it, a transient shift of the picks leads to a propagation of the signal. As indicated by arrows, the first column of cells goes up, followed by the second, the third one and so on.

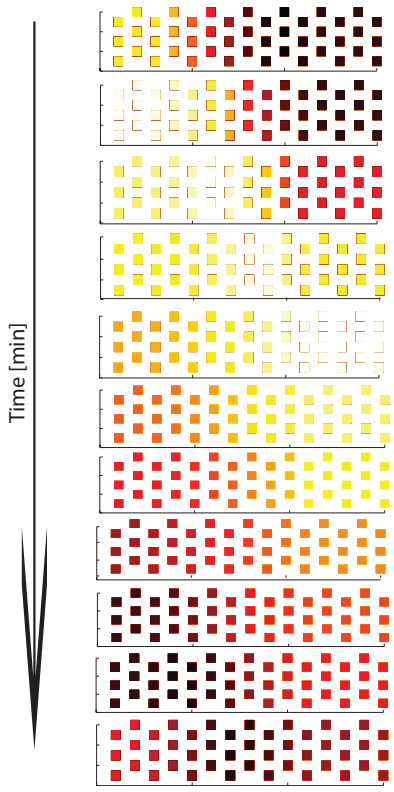


Figure 9.3: **Propagation along a stripe, in space.** Simulation on a 3×15 lattice of cells of a pulse propagation. Cells are modeled as autonomous oscillators and they are able to communicate with the 6 immediately neighbouring cells. The parameters are chosen to obtain a region of synchronization so that a pulse given on the boundary conditions on first column of cells is able to propagate across the tissue. White signal represent a low level of expression, while black one corresponds to a high level.

Propagation can not be obtained in the case of an unsynchronized state,

thus being a phenomenon possible only in the case an "ordered system state" as in the case of synchronization.

9.3.1 Non-oscillating PSM cells scenario

Here, we move to consider the simple case in which PSM cells do not oscillate except for a small group at the tip of the elongating embryo.

To simulate this kind of situation, I set the parameter α to 1, thus removing the negative autoregulatory feedback. Mathematically, the cells reach a steady state and they do not present oscillations, also in the case they are stimulated. When a short pulse is applied to the first column of cells, this is the first to react, followed by the second, the third and then all the others. Effectively, as shown in Figure 9.4 and Figure 9.5, the pulse propagates from one side to the other with a strong dampening. With the term "propagation", I mean the number of stripes that are involved in the signal transmission.

The same happens when increasing Notch and Delta production rates (moving b_D and b_N from $(40, 40)$ to $(60, 60)$).In Figure 9.6 a simulation over space and time is presented: the embryo is represented by a rectangle of dimensions 3×29. Each square is a cell and it is able to interact with the other six neighbouring cells. At time 0 the "tissue" presents an uniform expression of the reporter gene. When stimulated with a single pulse at the first column of cells, it reacts, losing this homogeneous expression: at the beginning, a spatial transient can be observed with the first columns of cells

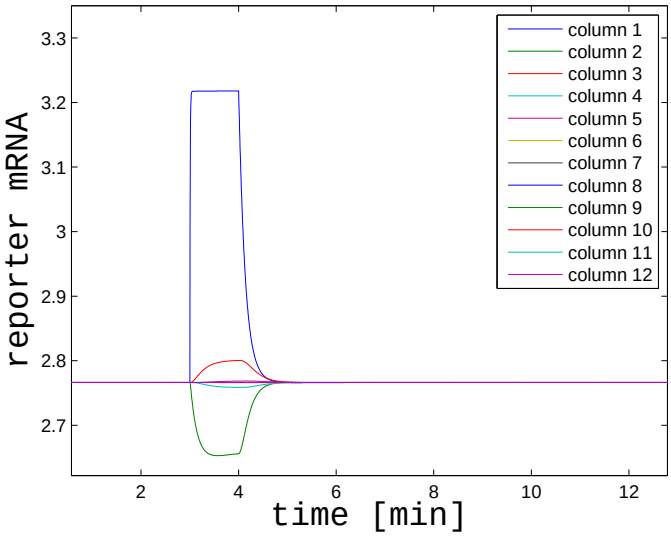


Figure 9.4: **Simulation over time of a single pulse on a tissue of non oscillating cells.** At time $t = 3$ a short pulse is applied to the first column of cells in a stripe: each line corresponds to a single cell but in different positions along the stripe (columns).

presenting opposite expression; after this phase, a gradient in expression can be observed for further columns. The final columns are too far to react to this perturbation. Over time this difference in expression is lost, in line to what is shown in Figure 9.4.

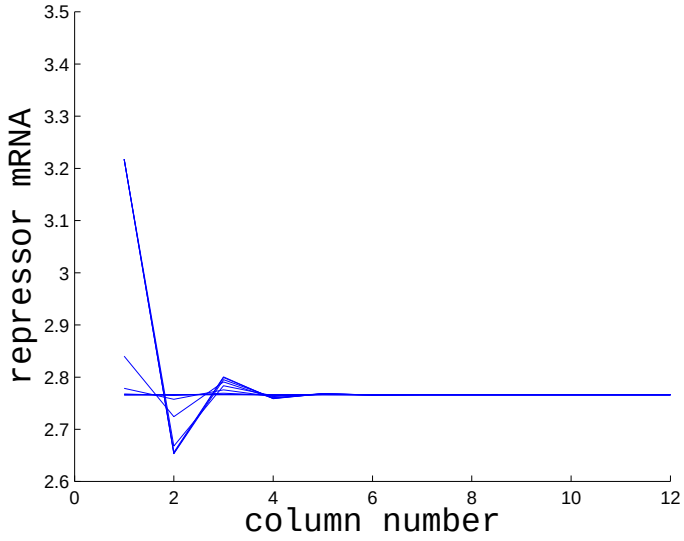


Figure 9.5: **Simulation in space of a single pulse on a tissue of non oscillating cells.** At time $t = 3$ a short pulse is given to the first column of cells in a stripe: each line corresponds to a different time instant.

9.3.2 PSM cells as dampened oscillators

I then considered cells as dampened oscillators, thus not able to oscillate autonomously but only when perturbed. This is a plausible assumption since PSM cells cultured in vitro exhibit weak oscillations with great variability across the single cells [88]. To this end, I set the parameter α to 0.9.

Differently from the case of non-oscillating cells, in which the profile clearly reveals a dampening of the stimulus across the tissue, in this case the cells present an oscillatory expression, with their height dampened over

time. So, a wave effectively propagates across the cells as can be seen from Figure 9.7. Here, we can distinguish between temporal dampening and spatial propagation: with "temporal dampening" I mean the oscillation's amplitude decay over time, while by "spatial propagation", I mean how far across the stripe the wave moves. An increment in b_D and b_N decreases dampening but does not affect spatial propagation. For this kind of system, we gave a single pulse as in the case described in Section 9.3.1. Also in this case, a spatial transient phase is observed but, for after this, a wave propagation can be observed, as shown in Figure 9.7. The wave travels with a dampening in space (so that further cells cannot sense the signal) and over time (so that the wavefront retrocedes in time). Thus, differently from the case in which the negative feedback is completely stopped, wave propagation is possible for a network of dampened oscillatory cells. Moreover, in order to simulate the action of a group of pacemaker cells, a periodic external pulse can be given at the tip of the embryo: results are shown in Figure 9.8, where it is possible to see that wave propagation is always possible and the oscillation can stop after a fixed number of cells' column without the action of an external morphogen. This ability of the network to determine autonomously where the wave have to stop is thus entirely dependent on the properties of the single cells, as seen comparing the behaviour of the system with that one described in 9.3.

9.4 Conclusion

In conclusion, I showed that it is possible to have a wave propagation even when cells are not autonomous oscillators and that, in this case, no gradient is needed to stop the wave propagation along the stripe, comparing Figure 9.3 to Figure 9.8. I suggest that this mechanism can be used to modify the "Clock and Wavefront" model and yield to a new one in which the cells are dampened clocks. Here, the gradient is not required in order to stop the wave at the right place, but this is an intrinsic property of the tissue. This mechanism is analogous to a ball rolling along a plane: without friction, it will never stop, but adding it to the surface (i.e. dampening the oscillations) the ball can stop. This traveling wave, provides the timing during the somite formation and could also contribute to the formation of rostro-caudal polarity [43].

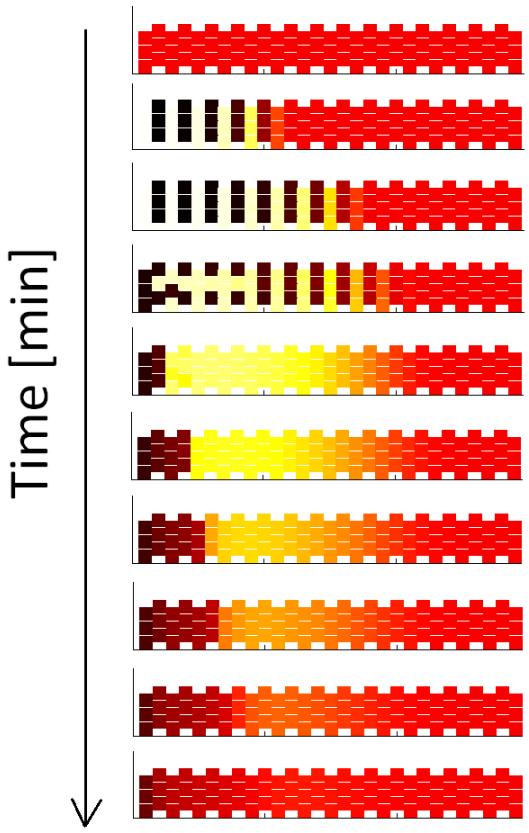


Figure 9.6: **Propagation along a stripe, in space, of a single pulse when cells are non-oscillating cells.** Simulation on a 3×29 lattice of cells of a wave propagation. Cells are modeled at the stable steady state and they are able to communicate with the 6 immediately neighbouring cells. White signal represent a low level of expression, while black one corresponds to a high level. Starting from a uniform distribution, the cells sense the signal one column at time with a strong dampening effect.

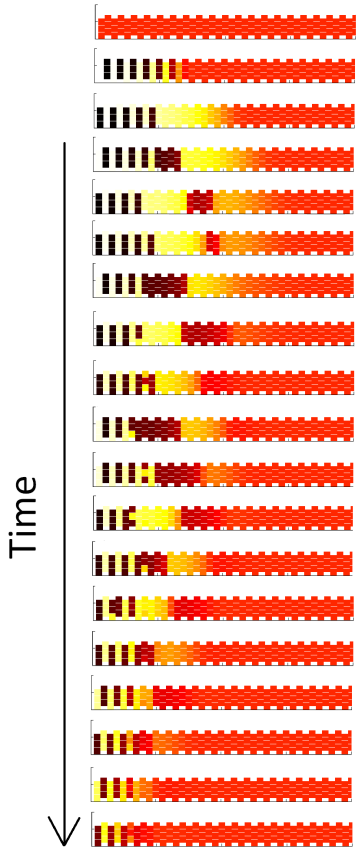


Figure 9.7: **Propagation along a stripe, in space, when cells are damped oscillators.** Simulation on a 3×40 lattice of cells of a pulse propagation. Cells are modeled as dampened oscillators at the steady state. A pulse given on the first column of cells is able to propagate across the tissue. White signal represent a low level of expression, while black one corresponds to a high level. Starting from a uniform distribution, the cells are able to oscillate: after a spatial transient of some columns, in which cells oscillate in antiphase, the propagation occurs along the stripe until the signal stops when too far from its source. Over time the "wavefront" retrocedes up to the first lines, determing at every time the position for a new couple of somites.

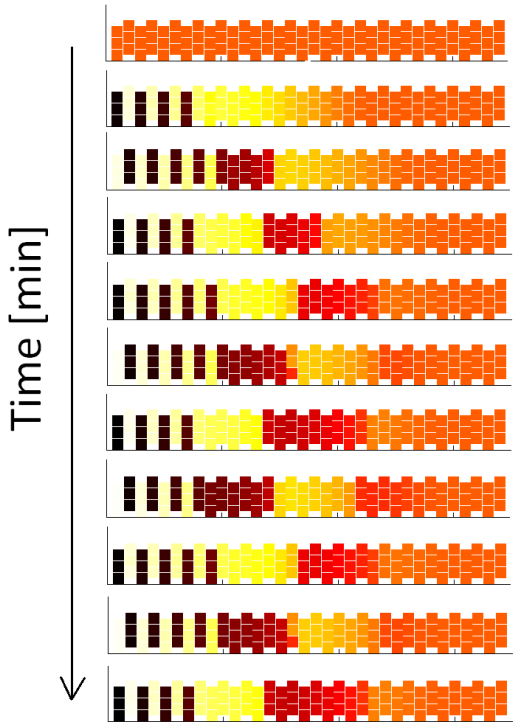


Figure 9.8: **Propagation along a stripe, in space, when cells are periodically entrained dampened oscillators.** Simulation on a 3×30 lattice of cells entrained by an external periodic input. Cells are modeled as dampened oscillators, initially at the steady state and they are able to communicate with the 6 immediately neighbouring cells. White signal represent a low level of expression, while black one corresponds to a high level. Starting from an uniform distribution, the cells are able to oscillate in a sustained manner and to propagate the signal coming from the external input. This signal is dampened over space, thus determining the possible wavefront where the oscillations can stop.

Chapter 10

Conclusion

In this Thesis, modeling and analysis of synthetic networks and endogenous systems have been discussed at single cell and tissue level. After an overview on the tecniques used to mathematically describe biological systems, the comparison between different network topologies is perfomed: the circuits considered are synthetic in order to exactly know what are the interacting components. These networks have been integrated in a mammalian cell line. Beside the role of the different feedbacks involved, also the role of biological parts such as the microRNA in biological systems has been investigated. All these analyses are reported in Chapters 3, 4, 5.

Then, after the analysis of synthetic circuits, this thesis considered endogeneous systems. Moreover, starting from the analyses obtained in the previous chapters, in Chapter 9 a simulation and a new simple mechanism is provided to simulate and explain a developmental process common to all

vertebrates: somite segmentation. Somites are supposed to be self-organizing structures independent of the oscillating activity of Notch-Delta genes. I proposed a model which is able to reproduce the behaviour and the main features of traveling transcriptional wave observed *in vivo*, without the further complication of an external morphogen.

Chapter 11

Appendix A

Materials and Methods for construction and treatment of the PFL

Experimental procedure: determination of d2EYFP half-life

To evaluate d2EYFP degradation rate, 500 stably integrated CHO tetOFF cells were plated in chamber slide (lab-Tek) and, after cell adhesion, Cycloheximide (Sigma, stock dilution 10 mg/ml in sterile water) was added to the medium to a final concentration of 10μg/mL, 50μg/mL, 100μg/mL or 500 500μg/mL. Temperature was maintained at 32 °C. Image acquisition and analysis was performed as described above. The experimental data were fitted

into an exponential curve using the curve fitting tool (cftool) from Matlab 2010b, and the degradation coefficient d_3 was used to obtain the half-life (τ) of the d2EYFP protein: τ=log(2)/ d_3

Methods for the PNFL costruction and analysis

Switch-OFF and switch-ON experiments

PFL7 and PNFL 7.2 cells were plated in a 12-well culture plate in order to obtain 400,000 cells per well, and then treated (or mock-treated) with 1 μg/ml Doxycycline. Fluorescence was measured by Fluorescence-Activated Cell Sorting (FACS) analysis at 0, 6, 12, 24, 48 and 72 hrs after treatment. The Doxycycline-containing medium was replaced 48 hrs after treatment. For the switch-ON experiments, PFL7 and PNFL 7.2 cells were pre-treated with 1 μg/ml Doxycycline for 72 hrs and then plated as described above, either in medium containing Doxycycline or not (control). Fluorescence was measured by FACS at 0, 24, 48, 72 and 96 hrs after treatment. For both switch-OFF and switch-ON experiments, each FACS measurement was made on a sample of 20.000 events in duplicate.

Microfluidics and Microscopy

We used the microfluidic device for mammalian cell culture described in [76]. Before cell loading, the channels in the device were coated with 100 μg/ml fibronectin for 1 hour at room temperature. After coating, the fibronectin solution was removed using a vacuum manifold, and a high density cell suspension (10^7 cells/ml) was loaded in the dedicated inlet. PFL and PNFL cells were loaded into the cell traps at a seeding density of 10 to 20 cells per trap, and the device was placed in a cell culture incubator to allow attachment of cells. The device was then placed under a Nikon Eclipse TI-E inverted epifluorescence microscope, equipped with an incubation chamber (H201-OP R2, Okolab). The two syringes connected to the inlets were filled with untreated medium, or medium containing 1 μg/ml doxycycline. Sulforhodamine 101 (Sigma-Aldrich, Germany), a fluorescent dye, at a concentration of 1 μM was added to the medium containing doxycycline. The flow was adjusted to provide doxycycline for a given time interval (40 min, 60 min, 180 min for PNFL cells and 960 min or 1800 min for PFL cells). Bright field and fluorescence images of 3 to 20 traps per experiment were acquired with a 40x objective at intervals of 15 min. Fluorescence images were taken using FITC (for d2EYFP, excitation 460/40 nm, emission 510/50 nm) and TRITC (for Sulforhodamine 101: excitation 530/30 nm, emission 590/60 nm) filters. Images were acquired using a Peltier-cooled Andor Clara camera controlled by Nikon Instrument Software AR v.3.22.14.

Doxycycline treatment of PFL and PNFL clones with the microfludics device

PFL and PNFL cells were loaded into the cell traps of the microfluidic device at a seeding density of 10 to 20 cells per trap; the average fluorescence intensity of the cells in at least 3, and up to 20, traps per experiment was tracked by time-lapse microscopy at 15 min intervals. The microfluidics platform enables to change in real-time growth conditions of cells by switching between two different media. We provided the cells with untreated growth medium for two hours, to rule out any fluctuation in metabolism due to displacement of the device from the cell incubator to the incubation chamber of the microscope; we then switched to Doxycycline-treated medium (1 $\mu g/mL$) for different time intervals, according to the simulations: 960 min (Figure 2C in the main text) or 1800 min (Figure 2D in the main text) for PFL cells, and 60 min (Figure 3C in the main text) or 240 min (Figure 3D in the main text) for PNFL cells. After the Doxycycline pulse, cells were switched again to untreated medium for the rest of the experiment. We quantified the mean fluorescence for each group of cells in a trap, as to rule out cell-to-cell variability, and we plotted the mean and standard deviation among all traps (i.e. replicates) for each experiment.

Image acquisition and processing

Images were acquired using an inverted epifluorescence microscope (Nikon Eclipse TI-E, Nikon Instruments) equipped with an incubation chamber (H201-OP R2 ,Okolab), a digital camera (Andor Clara, Andor), a 20X objective (Obj. CFI PF DLL 20X Ph1, Nikon) and a 512-nm/529-nm (B/G/R) d2EYFP-specific excitation/emission filter set. Temperature was maintained at a constant level as the experimental setup required, while CO2 concentration was set to be 5% of the total air volume injected in the incubation chamber. Both phase-contrast images and fluorescent fields were acquired at intervals of 15 minutes. Exposure times for the phase-contrast field was set to 2 ms (transmitted light lamp voltage was set to 4.5 V) while 300 ms (Intensilight lamp set at 10% of the maximum power) was chosen as exposure time for the fluorescent images: this choice was meant to prevent photobleaching while optimising the ratio between the quality of the images and reflected-light-induced stress on the cells. Experiments were carried out using NIS-Elements AR v.3.10 644 (Nikon Instruments) software package and the Perfect Focus System (Nikon Instruments) to maintain the same focal plane during the whole duration of the experiment. At the end of the acquisition process, images were extracted as raw data for the fluorescence quantification procedures.

The experiments were set up so that at the beginning of each experiment the first image contained at least 15 cells and no more than 30 cells, to avoid

cells exiting the image during the time-lapse experiment due to cell replication and "over-crowding". Image segmentation was carried out in Mathworks Matlab R2010b (Mathworks Inc.); the algorithm we implemented to quantify fluorescence was meant to distinguish the foreground (living cells) from the background in each image of the bright field. We used morphological operators such as erosion and dilation (*imerode* and *imdilate* functions from the MATLAB image processing toolbox). Thus two binary masks were built in order to compute separately the mean d2EYFP fluorescence of the foreground and the background using an element by element matrix multiplication between the binary images and the fluorescent one. The average fluorescence intensity across the cell population was then computed as the difference between the foreground and the background for each image at each time point (i.e. no single cell fluorescence quantification is performed).

Model simulations and parameter identification

Numerical simulations were run using Matlab 2010b (Mathworks Inc.). We used *ode23s* solver (a detailed discussion of the numerical methods used by *ode23* can be found in [22]). For the parameter identification, we used the PottersWheel toolbox [84] implemented in MATLAB. Two sets of parameters were identified: the dynamical parameters governing the model and a scaling

factor meant to approximate the transduction contribution of the microscopy equipment. Since Doxycycline has been only added at time $t = 0$ min in our experiment we forced the fitting procedures to start from the model predicted ON steady state.

We defined the following objective function:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_{Exp}(i) - y_{Model}(i))^2}{\sigma_{Exp}(i)^2} \tag{11.1}$$

where N is the number of experimental data points, y_{Model} are the predicted values of the mathematical model (using the inferred parameters), y_{exp} are the experimental data points and $\sigma_{Exp}(i)^2$ is the sample variance computed over the experimental replicates.

As optimisation algorithm we used Trust Region Method (TRM) in a logarithmic parameter space: at the k^{th} iteration of the optimisation procedure the TRM approximates the shape of the function f to be minimised with the model m_k thus trying to solve the following problem:

$$\min_{p} m_k(x_k + p_k) \tag{11.2}$$

being $x_k + p_k$ the new parameter vector considered as solution at the k^{th} iteration. If the model m_k has quadratic form the vector p_k can be obtained by observing that:

$$m_k(x_k + p_k) = f_k + p_k^T + \nabla f_k + \frac{1}{2} p_k^T B_k p_k \tag{11.3}$$

and therefore

$$p_k = -\nabla_k \frac{\nabla f_k}{\| \nabla f_k \|} \tag{11.4}$$

In order to allow an extensive exploration of the parameters' space, and to avoid local minima, we used a quasi-random number generator routine in PottersWheel [84] to select an initial guess of the parameters' values, and then launched the TRM procedure M times (M=100 in our settings), requiring the cost function in eq. 11.1 to be $\chi^2/N \leq 0.5$ [84].

The values in Table 1 represent parameters for which the cost function (eq. 11.1) is the smallest across the M runs; whereas the standard devation of each parameter in Table 1 is evaluated by considering all of the M runs.

Moreover, in order to compare switch off times among the different experiments, we computed the τ_{off} defined as the time the circuit needed to achieve the 50% of the mean initial fluorescence MIF calculated for each experiment as follows:

$$MIF = \frac{\sum_{i=1}^{3} f_i}{3} \tag{11.5}$$

with f_i fluorescence of the i^{th} frame in the sequence smoothed by moving average filtering.

11.1 Algorithms for network simulations

Here, we provide some matlab codes useful to perform patterning simulatons on a network of cells with periodic boundary conditions and the modified codes to obtain the simulation of wave propagation along a stripe of cells (fixed boundary conditions).

Main

Here, the main program to simulate the behaviour of a network of cells. The LI model is considered in this case.

```
clear all
pack
clc

I = 30; %rows
M = 3; %number of cells for every row

bD = 40; %set bD
bN = 40; %set bN
k = 1;
```

```
tend = 300; \%set the final time for integration
M = M*I; \%total cell number


par(1) = bN;      %bN
par(2) = 1;       %gamma
par(3) = 1;       %kt
par(4) = 1;       %m
par(5) = 24300; %bR
par(6) = 1;       %n
par(7) = 300000;  %kRS
par(8) = 1;       %gammaR
par(9) = bD;     %bD
par(10) = par(9); %bD
par(11) = 0.55;   %kc
par(12) = 0.1;    %p0
par(13) = 33;    %K
par(14) = 0.5;  %aH
par(15) = 1; %bP
par(16) = 0.135; %Kf
par(17) = 0.01; %Kp


[t y] = ode23s(@LI_delayHSS,[0 1000],[10 10 10 10 10],[],[bN bD],0);
```

```
Not = y(end,1);

Dll = y(end,2);

R = y(end,3);

p1 = y(end,4);

p2 = y(end,5);


INIT1 =  [Not Dll R p1 p2];

INIT = INIT1;


for i=1:M/I-1

INIT= [INIT;Not Dll R p1 p2];

end


%solving the homogeneous steady state

[t y] = ode23s(@LI_delayHSS,[0 1000],[10 10 10 10 10],[],[bN bD],0);

Not = y(end,1);

Dll = y(end,2);

R = y(end,3);

p1 = y(end,4);

p2 = y(end,5);
```

```
%connection matrix construction:
%fixed boundary condition case
%periodic boundary condition case

L = matconn_retesag_NOPeriodic(I,M/I);

for i=M/I:M-1
 INIT= [INIT;Not Dll R p1 p2];
end

int=0.01; %doppio del passo di integrazione
s=1;
Not = Not;%+random('Uniform',-Not*s,Not*s);
Dll = Dll;%+random('Uniform',-Dll*s,Dll*s);
R = R;%+random('Uniform',-R*s,R*s);
p1 = p1;%+random('Uniform',-p1*s,p1*s);
p2 =  p2;%+random('Uniform',-p2*s,p2*s);

%setting the boundary condition:
%fixed boundary condition case
%periodic boundary condition case
BC = boundary_condition_striscia_pulse(I,M./I,par,Not,Dll,R,p1,p2,0);
```

```
%BC = boundary_condition_sp(I,M./I,par,Not,Dll,R,p1,p2,0);



%number of equation for each cell
N = 5;

%weight and adiancent matrix construction
for i=1:M
    D(i) = L(i,i);
end

K = diag(D);
G = - (L - K);

%solving the system of  M x N equations using 4 order Runge-Kutta method

tempi = [0:int:tend];

x0 = [];

options =[];
res1 = [];
```

```
res2 = [];

res3 = [];

res4 = [];

res5 = [];


for i=1:M

 x0 = [x0;INIT(i,:)'];


  res1 = [res1,INIT(i,1)]; %Ni

  res2 = [res2,INIT(i,2)]; %Di

  res3 = [res3,INIT(i,3)]; %Ri

  res4 = [res4,INIT(i,4)]; %p1i

  res5 = [res5,INIT(i,5)]; %p2i


end


v = (kron(inv(K),eye(N)))*((kron(G,eye(N)))*x0)+BC;

step = 2;

time = 0;


i = 1;
```

```
while (time(i)<=tend)

   res = [];
   BC = boundary_condition_s(I,M./I,par,Not,Dll,R,p1,p2,time(i)+int);
   AC(i,:)=BC;
   for j=1:M

     input = v(1+N*(j-1):N+N*(j-1));
     init = x0(1+N*(j-1):N+N*(j-1));

  [t y] = rk4fixed(@LI_delay,[time(i),time(i)+int],init,step,input,[bN bD]);

 res1(i+1,j) = [y(2,1)];
 res2(i+1,j) = [y(2,2)];
 res3(i+1,j) = [y(2,3)];
 res4(i+1,j) = [y(2,4)];
 res5(i+1,j) = [y(2,5)];

 res = [res;res1(i+1,j);res2(i+1,j);res3(i+1,j);res4(i+1,j);res5(i+1,j)];
 end

 time(i+1) = t(2);
```

```
 v = (kron(inv(K),eye(N)))*((kron(G,eye(N)))*res)+BC;
 i = i+1;
 x0 = res;

end

for i=1:M
    ss(i,:) = [res1(end,i) res2(end,i) res3(end,i) res4(end,i) res5(end,i)];
%ith cell
%figure,plot(time,res1(:,i),'b',time,res2(:,i),'r',time,res3(:,i),'g')
end
```

LI model

LI model matlab script useful to solve the system of $5 \times \mathcal{N}$ equations in a network of cells.

```
%LI model
function dy = LI_delay(t,y,input,par)

bN = par(1);%10;
bD = par(2);%100;
```

```
bR = 24300;%1000000;
Krs = 300000;
n=1;
alfa = 0.9;
%alfa=alfa*(1-(sign(t-3)+1)/2) + 0.8*((sign(t-3)+1)/2);
k = 33;
p0 = 0.1;
Kt= 1;
bP = 1;
Kf = 0.135;
Kp = 0.01;
g=1;
g1=1.;
kc=0.;
tau=1;
U1 = input(1); % Notchext;
 U2 = input(2); % Dllext;
%Notch
 dy(1,1) = tau*(bN - g*y(1) - (U2*y(1))*Kt -kc*y(1)*y(2));

%Dll1
 dy(2,1) = tau*(bD*(1./(1+y(5).^2))-(U1*y(2))*Kt-g1*y(2) -kc*y(1)*y(2));
```

```
%Reporter mRNA
 dy(3,1) = tau*((bR*(((y(1)*U2).^n)./(Krs + (y(1)*U2).^n))...
...)*(alfa + (1-alfa)*(k*(p0^2)./(p0^2+y(5)^2))) - y(3));

%reporter unfolded protein
 dy(4,1) = tau*(bP*y(3)-(1+Kf)*y(4));

%reporter folded protein
 dy(5,1) = tau*(Kf*y(4)- y(5)./(Kp + y(5)));
end
```

LI Model Homogeneous state

LI model rewritten to estimate the homogeneous steady state.

```
function dy = LI_delayHSS(t,y,par,alfa)

bN = par(1);%10;
bD = par(2);%100;
bR = 24300;%1000000;
Krs = 300000;
```

```
n=1;

alfa = 0.9;

k = 33;

p0 = 0.1;

kc=0.;

Kt= 1;

bP = 1;

Kf = 0.135;

Kp = 0.01;

g=1;

g1=1.;

tau=1;

%Notch

 dy(1,1) =tau*( bN - g*y(1) - (y(2)*y(1))*Kt -kc*y(1)*y(2));


%Dll1

 dy(2,1) = tau*(bD*(1./(1+y(5).^2))-(y(1)*y(2))*Kt- g1*y(2) -kc*y(1)*y(2));


%Reporter mRNA

 dy(3,1) = tau*((bR*(((y(1)*y(2)).^n)./(Krs + (y(1)*y(2)).^n))...

...)*(alfa + (1-alfa)*(k*(p0^2)./(p0^2+y(5)^2))) - y(3));
```

```
%reporter unfolded protein
 dy(4,1) = tau*(bP*y(3)-(1+Kf)*y(4));

%reporter folded protein
 dy(5,1) = tau*(Kf*y(4)- y(5)./(Kp + y(5)));
end
```

Connection Matrix

Here, two MATLAB script are reported to determine the Connection Matrix in the case of a periodic or non-periodic network. The network considered is rectangular and each cell is hexagonal, able to communicate with the others 6 neighbouring cells.

Non-periodic grid

```
%matrix  L construction

% J number of cells for each row
% I number of rows

function L = matconn_retesag_NOPeriodic(I,J)
```

```
A = 4*eye(J);
A(1,1) = 2;
A(J,J) = 3;
A = A -diag(ones(J-1,1),1)-diag(ones(J-1,1),-1);

F = A;
F(1,1)=3;
F(J,J)=2;

B=-1*eye(J)-diag(ones(J-1,1),-1);

D=-1*eye(J)-diag(ones(J-1,1),1);

%case of two rows
if (I<3)
    L = [A B; B A];
else

C = 6*eye(J);
C = C -diag(ones(J-1,1),1)-diag(ones(J-1,1),-1);
```

```
C1 =C;
C1(1,1) = 5;
C1(J,J) = 3;

C2 =C;
C2(1,1) = 3;
C2(J,J) = 5;

L = [A B];

for j=3:I;

    L = [L zeros(J)];

end

comp1 = [D C1 D];
comp2 = [B C2 B];

S = I - 3;
k = 0;
comp = comp1;
```

```
for i = 2:I-1
    riga = comp1;

    for t = 1:k

        riga = [zeros(J) riga];
    end

    for tt = 1:S
     riga = [riga zeros(J)];
    end
    S=S-1;
    k=k+1;

    L = [L;riga];

    var = comp1;
    comp1 = comp2;
    comp2 = var;

end
```

```
ultima_riga = [];


for j = 1:I-2
    ultima_riga = [ultima_riga zeros(J)];


end


if (mod(I,2)) %odd
ultima_riga = [ultima_riga B A];
else
ultima_riga = [ultima_riga D F];
end


L = [L ; ultima_riga];
end
end
```

Periodic grid

```
% J number of cells for each row
% I number of rows
```

```
function L = matconn_retesag_torus(I,J)


A = 6*eye(J);

A(1,J) = -1;

A(J,1) = -1;

A = A -diag(ones(J-1,1),1)-diag(ones(J-1,1),-1);


B=-1*eye(J)-diag(ones(J-1,1),-1);

B(1,J)=-1;


D=-1*eye(J)-diag(ones(J-1,1),1);

D(J,1)=-1;


%case of two rows

if (I<3)

    L = [A B; B A];

else


C = 6*eye(J);

C = C -diag(ones(J-1,1),1)-diag(ones(J-1,1),-1);


C(1,J) = -1;
```

```
C(J,1) = -1;


L = [A B];


for j=3:I-1;


    L = [L zeros(J)];


end
L = [L B];
comp1 = [D C D];
comp2 = [B C B];


S = I - 3;
k = 0;
comp = comp1;


for i = 2:I-1
    riga = comp1;


    for t = 1:k
```

```
        riga = [zeros(J) riga];
    end

    for tt = 1:S
     riga = [riga zeros(J)];
    end
    S=S-1;
    k=k+1;

    L = [L;riga];

    var = comp1;
    comp1 = comp2;
    comp2 = var;

end

ultima_riga = [D];

for j = 1:I-3
    ultima_riga = [ultima_riga zeros(J)];
```

```
end

if (mod(I,2)) %odd
ultima_riga = [ultima_riga B A];
else
ultima_riga = [ultima_riga D A];
end

L = [L ; ultima_riga];
end
end
```

Fixed Boundary Conditions

MATLAB script to set the boundary conditions for a network of cells. The network is considered to be rectangular and each cell an hexagon, able to communicate with the other 6neighbouring cells. In this case, boundary conditions are considered fixed to zero around the network and a non-zero value is consider only for one side of the rectangular network.

```
function BC = boundary_condition_s(I,J,par,Not,Dll,R,p1,p2,time)

BC = [];
```

```
options =[];
input = [0 0];

bN=par(1);
bD = par(9);

%D=25*((((sign(time-2.74)+1)/2-(sign(time-28)+1)/2)));

NN_1 = Not*D;
DD_1 = Dll*D;
RR_1 = R*D;
RR_p1 = p1*D;
RR_f1 = p2*D;

app_1 = [NN_1 DD_1 RR_1 RR_p1 RR_f1];
app_0 = [NN_1 DD_1 RR_1 RR_p1 RR_f1];

app = [app_0;app_1];

%first row
i = 1;
```

```
for j = 1:J

    if (j==1 || j==J)

     BC = [BC;[app(1,:)']];



    else
        BC = [BC;app(1,:)'];

    end

    var = app(1,:);
    app(1,:)=app(2,:);
    app(2,:)=var;
end

%middle rows

if (I>1)
    if (I>2)
```

```
for i=2:I

    for j=(i-1)*J+1:i*J
        if (j==(i-1)*J+1 || j==i*J)
          BC = [BC;[0 0 0 0 0]'];

          if (j==(i-1)*J+1)

          end
        else
            BC = [BC; [0 0 0 0 0]'];
        end

    var = app(1,:);
    app(1,:)=app(2,:);
    app(2,:)=var;
    end

end

end
```

```
end

end
```

Periodic Boundary Conditions

MATLAB script to set periodic boundary conditions for a network of cells.

```
function BC = boundary_condition_sp(I,J,par,Not,Dll,R,p1,p2,time)


BC = [];


options =[];

input = [0 0];


bN=par(1);

bD = par(9);


if time==0

[t y] = ode23s(@LI_delayHSS,[2,2+0.1],[Not Dll R p1 p2],options,[bN bD]);

NN_1 = y(1,1);

DD_1 = y(1,2);
```

```
RR_1 = y(1,3);

RR_p1 = y(1,4);

RR_f1 = y(1,5);


else

[t y] = ode23s(@LI_delayHSS,[2,time+2],[Not Dll R p1 p2],options,[bN bD]);

NN_1 = y(end,1);

DD_1 = y(end,2);

RR_1 = y(end,3);

RR_p1 = y(end,4);

RR_f1 = y(end,5);

end



if time==0

[t y] = ode23s(@LI_delayHSS,[2+0,2+0.1],[Not Dll R p1 p2],options,[bN bD]);

NN_0 = y(1,1);

DD_0 = y(1,2);

RR_0 = y(1,3);

RR_p0 = y(1,4);

RR_f0 = y(1,5);

else
```

```
[t y] = ode45(@LI_delayHSS,[2+0,2+time],[Not Dll R p1 p2],options,[bN bD]);

NN_0 = y(end,1);
DD_0 = y(end,2);
RR_0 = y(end,3);
RR_p0 = y(end,4);
RR_f0 = y(end,5);
end

app_1 = [NN_1 DD_1 RR_1 RR_p1 RR_f1];
app_0 = [NN_0 DD_0 RR_0 RR_p0 RR_f0];

app = [app_0;app_1];

%first row
i = 1;

for j = 1:J

    if (j==1 || j==J)

     BC = [BC;[app(1,:)']];
```

```
    else
        BC = [BC;app(1,:)'];
    end

    var = app(1,:);
    app(1,:)=app(2,:);
    app(2,:)=var;
end

%middle rows
if (I>1)
    if (I>2)

for i=2:I
    for j=(i-1)*J+1:i*J
        if (j==(i-1)*J+1 || j==i*J)
          BC = [BC;[0 0 0 0 0]'];

          if (j==(i-1)*J+1)

          end
```

```
        else
            BC = [BC; [0 0 0 0 0]'];
        end

    var = app(1,:);
    app(1,:)=app(2,:);
    app(2,:)=var;
    end

end

end

end
end
```

Runge-Kutta

Runge-Kutta MATLAB script to integrate the netwrok system equations.

```
function [T,X]=rk4fixed(Fcn,Tspan,X0,N,v,par)

% Matlab implementation of a fixed-step RK4 algorithm.
```

```
% Written by R.G. Longoria 950129.


h = (Tspan(2)-Tspan(1))/N;

halfh = 0.5*h;

%

neqs=size(X0);

X=zeros(neqs(1),N);

T=zeros(1,N);

X(:,1)=X0;

T(1)=Tspan(1);

Td = Tspan(1);

Xd = X0;

for i=2:N,

 [RK1]= feval(Fcn,Td,Xd,v,par);

 Thalf = Td + halfh;

 Xtemp = Xd + halfh*RK1;

 [RK2] = feval(Fcn,Thalf,Xtemp,v,par);

 Xtemp = Xd + halfh*RK2;

 [RK3] = feval(Fcn,Thalf,Xtemp,v,par);

 Tfull = Td + h;

 Xtemp = Xd + h*RK3;

 [RK4] = feval(Fcn,Tfull,Xtemp,v,par);
```

```
 X(:,i) = Xd + h*(RK1+2.0*(RK2+RK3)+RK4)/6;
 T(i) = Tfull;
 Xd = X(:,i);
 Td = T(i);
end
X=X';T=T';
```

Acknowledgements

Per finire, vorrei ringraziare tutti quelli che, di questi tre anni, hanno festeggiato l'inizio e ne festeggeranno la fine.

Il Dr. Diego di Bernardo, per aver dato il "via" a questo progetto, dandomi l'occasione di far parte del suo laboratorio; per aver capito e assecondato con fiducia le mie necessità ma, soprattutto, per avermi incoraggiato dinanzi ai numerosi tentativi fallimentari, che ho scoperto spesso necessari per realizzare un risultato importante.

Il Prof. Mario di Bernardo, per il suo aiuto scientifico, ma soprattutto per i suoi straordinari corsi che hanno stimolato la mia curiosità e fantasia: le migliori lezioni a cui avessi mai partecipato.

I would like also to thank Prof. John Hogan, for our interesting and illuminating discussions and for the opportunity he gave me to spend some weeks at the University of Bristol.

E, naturalmente, ringrazio tutti i ragazzi che dal 2010 ad oggi sono stati classificati in una delle sigle *gruppodibernardo*, *bioinformatici* e *auletta reietta*, amici più che colleghi: per le nostre sconclusionate discussioni a pranzo e le

numerosissime ore passate insieme, sia dentro che fuori il TIGEM; per aver allegerito anche i giorni più pesanti. Da ognuno di voi ho imparato qualcosa e sono certa che nulla di quello che è cominciato tra noi finirà.

Le mie allieve, per aver regalato alle mie serate qualcosa di così intenso che non avrei mai potuto immaginare.

In questi anni tutto quello che ho fatto, dalle lezioni e i congressi alle ore passate a studiare o a preparare le lezioni, è stato così fuori dalle aspettative che, a volte, mi sembrava di vivere una vita che non pensavo potesse essere la mia.

E, infine, ringrazio Annibale e la mia famiglia che, oltre a sopportare i miei lati peggiori quotidianamente, hanno saputo supportarmi più di tutti rendendo possibile tutto questo.

Bibliography

[1] Generic behavior of master-stability functions in coupled nonlinear dynamical systems. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 80(036204), Sep 2009.

[2] Allison L. Abbott, Ezequiel Alvarez-Saavedra, Eric A. Miska, Nelson C. Lau, David P. Bartel, H. Robert Horvitz, and Victor Ambros. The let-7 microrna family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in caenorhabditis elegans. *Developmental Cell*, 9:403 – 414, 2005.

[3] Murat Acar, Attila Becskei, and Alexander Van Oudenaarden. Enhancement of cellular memory by reducing stochastic transitions. *Nature*, 435:228–32, 2005.

[4] Bruce Alberts. *Molecular biology of the cell*. Garland Science, 2008.

[5] Kyle R. Allison and James J. Collins. Bacteria as control engineers. *Molecular Cell*, 41(1):4–5, 2011.

[6] Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall/CRC Mathematical & Computational Biology)*. Chapman and Hall/CRC, 2006.

[7] Ezequiel Alvarez-Saavedra and H. Robert Horvitz. Many families of c. elegans microrna s are not essential for development or viability. *Current Biology*, 20:367 – 373, 2010.

[8] Panos J. Antsaklis and Anthony N. Michel. *Linear Systems*. Birkhäuser Boston, corrected edition, October 2005.

[9] Alex Arenas, Albert Diaz-Guilera, Jurgen Kurths, Yamir Moreno, and Changsong Zhou. Synchronization in complex networks. *Physics Reports*, 469(3):93–153, December 2008.

[10] Raj Arjun and Alexander van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2):216 – 226, 2008.

[11] Spyros Artavanis-Tsakonas, Matthew D. Rand, and Robert J. Lake. Notch signaling: Cell fate control and signal integration in development. *Science*, 284(5415):770–776, April 1999.

[12] Alexander Aulehla, Winfried Wiegraebe, Valerie Baubet, Matthias B. Wahl, Chuxia Deng, Makoto Taketo, Mark Lewandoski, and Olivier

Pourquie. A [beta]-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nat Cell Biol*, 10.

[13] Daehyun Baek, Judit Villén, Chanseok Shin, Fernando D. Camargo, Steven P. Gygi, and David P. Bartel. The impact of microRNAs on protein output. *Nature*, 455(7209):64–71, 2008.

[14] Ruth E. Baker and Santiago Schnell. How can mathematics help us explore vertebrate segmentation? *HFSP Journal*, 3(1):1–5, 2009.

[15] Manuel Barrio, Kevin Burrage, André Leier, and Tianhai Tian. Oscillatory regulation of hes1: Discrete stochastic delay modelling and simulation. *PLoS computational biology*, 2(9), September 2006.

[16] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–593, June 2000.

[17] Attila Becskei, Bertrand Seraphin, and Luis Serrano. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J*, 20(10):2528–2535, May 2001.

[18] Lewin Benjamin. *Genes IX*. Jones and Bartlett, 2007.

[19] Matthew R Bennett and Jeff Hasty. Microfluidic devices for measuring gene network dynamics in single cells. *Nature reviews. Genetics*, 10:628–38, 2009.

[20] William J. Blake, Mads KAErn, Charles R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.

[21] Leonidas Bleris, Zhen Xie, David Glass, Asa Adadey, Eduardo Sontag, and Yaakov Benenson. Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular Systems Biology*, 7, 2011.

[22] P. Bogacki and L. F. Shampine. A 3(2) pair of runge-kutta formulas. *Appl. Numer. Math.*, 2:1–9, 1989.

[23] Sarah J. Bray. Notch signalling: three ways to use a pathway. *Cell & Developmental Biology*, 9(98):591–597, 1998.

[24] Sarah J. Bray. Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol*, 7(9):678–689, September 2006.

[25] Brian D Brown, Bernhard Gentner, Alessio Cantore, Silvia Colleoni, Mario Amendola, Anna Zingale, Alessia Baccarini, Giovanna Lazzari, Cesare Galli, and Luigi Naldini. Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. *Nat Biotech*, 25:1457–1467, 2007.

[26] Oriol Canela-Xandri, Francesc Sagus, Ramn Reigada, and Javier Buceta. A spatial toggle switch drives boundary formation in development. *Biophysical Journal*, 95(11):5111 – 5120, 2008.

[27] Irene Cantone, Lucia Marucci, Francesco Iorio, Maria Aurelia Ricci, Vincenzo Belcastro, Mukesh Bansal, Stefania Santini, Mario di Bernardo, Diego Di Bernardo, and Maria Pia Cosma. A yeast synthetic network for in vivo assessment of Reverse-Engineering and modeling approaches. *Cell*, 137(1):172–181, 2009.

[28] Richard W Carthew. Gene regulation by microRNAs. *Current Opinion in Genetics & Development*, 16:203–208, 2006.

[29] Xiao Chang, Zengrong Liu, Luonan Chen, and Ruiqi Wang. Bistability and oscillations in gene regulation mediated by small noncoding RNAs. *PLoS ONE*, 6:10, 2011.

[30] Angela H. Chau, Jessica M. Walter, Jaline Gerardin, Chao Tang, and Wendell A. Lim. Designing synthetic regulatory networks capable of Self-Organizing cell polarization. *Cell*, 151:320–332, 2012.

[31] Olivier Cinquin. Understanding the somitogenesis clock: Whats missing? *Mechanisms of Development*, 124:501 – 517, 2007.

[32] Joanne R. Collier, Nicholas A.M. Monk, Philip K. Maini, and Julian H. Lewis. Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. *Journal of Theoretical Biology*, 183(4):429 – 446, 1996.

[33] G Cuccato, G Della Gatta, and D di Bernardo. Systems and synthetic biology: tackling genetic networks and complex diseases. *Heredity*, 102(6):527–532, June 2009. PMID: 19259117.

[34] Giulia Cuccato, Athanasios Polynikis, Velia Siciliano, Mafalda Graziano, Mario di Bernardo, and Diego di Bernardo. Modeling RNA interference in mammalian cells. *BMC Systems Biology*, 5:1–9, 2011.

[35] Qinghua Cui, Zhenbao Yu, Enrico O Purisima, and Edwin Wang. MicroRNA regulation and interspecific variation of gene expression. *Trends in Genetics: TIG*, 23:372–375, 2007.

[36] Tal Danino, Octavio Mondragón-Palomino, Lev Tsimring, and Jeff Hasty. A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–330, January 2010.

[37] Eric H. Davidson. Emerging properties of animal gene regulatory networks. *Nature*, 468(7326):911–920, December 2010.

[38] Hidde De Jong. Modeling and simulation of genetic regulatory systems: A literature review. *J Comput Biol*, 9:67–103, 2002.

[39] T L Deans, C R Cantor, and J J Collins. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell*, 130(2):363 – 372, 2007.

[40] P. DeLellis, M.D. Bernardo, T.E. Gorochowski, and G. Russo. Synchronization and control of complex networks via contraction, adaptation and evolution. *Circuits and Systems Magazine, IEEE*, 10(3):64–82, thirdquarter 2010.

[41] Mary-Lee Dequeant and Olivier Pourquie. Segmental patterning of the vertebrate embryonic axis. *Nature Reviews Genetics*, 9(5):370–382, May 2008.

[42] Barbara Di Ventura, Caroline Lemerle, Kostantinos Michalodimitrakis, and Luis Serrano. From *in vivo* to *in silico* biology and back. *Nature*, 443:527–533, 2006.

[43] Ana S. Dias, Irene de Almeida, Julio M. Belmonte, James A. Glazier, and Claudio D. Stern. Somites without a clock. *Science*, page 1247575, January 2014.

[44] Margaret S Ebert and Phillip A Sharp. Roles for micrornas in conferring robustness to biological processes. *Cell*, 149:515–524, 2012.

[45] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000.

[46] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, Aug 2002.

[47] Abraham P. Fong, Zizhen Yao, Jun Wen W. Zhong, Yi Cao, Walter L. Ruzzo, Robert C. Gentleman, and Stephen J. Tapscott. Genetic and epigenetic determinants of neurogenesis and myogenesis. *Developmental cell*, 22(4):721–735, April 2012.

[48] Ursula Klingmuller Friedrich Marks, Karin Muller-Decker. *Cellular Signal Processing: An Introduction to the Molecular Mechanisms of Signal Transduction*. Garland Publishing.

[49] S. Pigolotti M.H. Jensen G. Tiana, S. Krishna and K. Sneppen. Oscillations and temporal signalling in cells. *Physical Biology*, (4):R1–R17, 2007.

[50] Jeffrey Gagan, Bijan K. Dey, and Anindya Dutta. MicroRNAs regulate and provide robustness to the myogenic transcriptional network. *Current Opinion in Pharmacology*, March 2012.

[51] T S Gardner, C R Cantor, and J J Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403:339–42, 2000.

[52] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342, 2000.

[53] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Biological Cybernetics*, 12(1):30–39, December 1972.

[54] François Giudicelli, Ertuğrul M Özbudak, Gavin J Wright, and Julian Lewis. Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS biology*, 5(6):e150, 2007.

[55] Leon Glass and Michael C. Mackey. *From Clocks to Chaos - The Rhythms of Life*. Princeton University Press, 1979.

[56] Albert Goldbeter and Olivier Pourquié. Modeling the segmentation clock as a network of coupled oscillations in the notch, wnt and {FGF} signaling pathways. *Journal of Theoretical Biology*, 252(3):574 – 585, 2008.

[57] B. Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3:425–437, 1965.

[58] M. Gossen and H. Bujard. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *PNAS*, 89:5547–5551, 1992.

[59] Raúl Guantes and Juan F. Poyatos. Dynamical principles of Two-Component genetic oscillators. *PLoS Computational Biology*, 2(3), March 2006.

[60] Jeff Hasty, David McMillen, and J J Collins. Engineered gene circuits. *Nature*, 420(6912):224–230, Nov 2002.

[61] Jeff Hasty, David McMillen, Farren Isaacs, and James J. Collins.

[62] Héctor Herranz and Stephen M. Cohen. MicroRNAs and gene regulatory networks: managing the impact of noise in biological systems. *Genes & Development*, 24:1339–1344, 2010.

[63] Frances A A. High and Jonathan A A. Epstein. The multifaceted role of notch in cardiac development and disease. *Nat Rev Genet*, December 2007.

[64] H. Hirata, S. Yoshiura, T. Ohtsuka, Y. Bessho, T. Harada, K. Yoshikawa, and R. Kageyama. Oscillatory expression of the bHLH factor hes1 regulated by a negative feedback loop. *Science (New York, N.Y.)*, 298(5594):840–843, October 2002.

[65] Hiromi Hirata, Yasumasa Bessho, Hiroshi Kokubu, Yoshito Masamizu, Shuichi Yamada, Julian Lewis, and Ryoichiro Kageyama. Instability of hes7 protein is crucial for the somite segmentation clock. *Nature Genetics*, 36(7):750–754, May 2004.

[66] E. Hornstein and N. Shomron. Canalization of development by microRNAs. *Nat Genet*, 38 Suppl 1, 2006.

[67] Itaru Imayoshi, Akihiro Isomura, Yukiko Harima, Kyogo Kawaguchi, Hiroshi Kori, Hitoshi Miyachi, Takahiro Fujiwara, Fumiyoshi Ishidate, and Ryoichiro Kageyama. Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science*, 342(6163):1203–1208, December 2013.

[68] Masafumi Inui, Graziano Martello, and Stefano Piccolo. Microrna control of signal transduction. *Nat Rev Mol Cell Biol*, 11:264–275, 2010.

[69] Farren J. Isaacs, Jeff Hasty, Charles R. Cantor, and J. J. Collins. Prediction and measurement of an autoregulatory genetic module. *Proceedings of the National Academy of Sciences*, 100(13):7714–7719, 2003.

[70] Mads Kaern, Timothy C. Elston, William J. Blake, and James J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6(6):451–464, 2005.

[71] Ryoichiro Kageyama, Yasutaka Niwa, Akihiro Isomura, Aitor Gonzlez, and Yukiko Harima. Oscillatory gene expression and somitogenesis. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1(5):629–641, 2012.

[72] Ahmad S. Khalil and James J. Collins. Synthetic biology: applications come of age. *Nature Reviews Genetics*, 11(5):367–379, May 2010.

[73] V. Narry Kim. MicroRNA biogenesis: coordinated cropping and dicing. *Nature Reviews Molecular Cell Biology*, 6:376–385, 2005.

[74] Hiroaki Kitano. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, March 2002.

[75] Hiroaki Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, November 2004.

[76] Martin Kolnik, Lev S Tsimring, and Jeff Hasty. Vacuum-assisted cell loading enables shear-free mammalian microfluidic culture. *Lab on a chip*, 2012.

[77] Beat P. Kramer, Alessando U. Viretta, Marie Daoud-El-Baba, Dominique Aubel, Wilfried Weber, and Martin Fussenegger. An engineered epigenetic transgene switch in mammalian cells. *Nat Biotechnol*, 22(7):867–870, July 2004.

[78] Beat P Kramer, Alessandro Usseglio Viretta, Marie Daoud-El Baba, Dominique Aubel, Wilfried Weber, and Martin Fussenegger. An engineered epigenetic transgene switch in mammalian cells. *Nat Biotechnol*, 22:867–870, 2004.

[79] Yan Kung, Weerawat Runguphan, and Jay D. Keasling. From fields to fuels: recent advances in the microbial production of biofuels. *ACS synthetic biology*, 1(11):498–513, November 2012.

[80] J. Lewis. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Current biology : CB*, 13(16):1398–1408, August 2003.

[81] Xin Li, Justin J Cassidy, Catherine A Reinke, Stephen Fischboeck, and Richard W Ca. A microrna imparts robustness against environmental fluctuation during development. *Cell*, 137:273–282, 2009.

[82] Yichen Li, Yumin Li, Hui Zhang, and Yong Chen. MicroRNA-Mediated positive feedback loop and optimized bistable switch in a cancer network involving miR-17-92. *PLoS ONE*, 6:e26302, 2011.

[83] Yusuke T. Maeda and Masaki Sano. Regulatory dynamics of synthetic gene networks with positive feedback. *Journal of Molecular Biology*, 359:1107 – 1124, 2006.

[84] Thomas Maiwald and Jens Timmer. Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, 24(18):2037–2043, September 2008. PMID: 18614583 PMCID: 2530888.

[85] Thomas Maiwald and Jens Timmer. Dynamical modeling and multi-experiment fitting with potterswheel. *Bioinformatics*, 24(18):2037–2043, 2008.

[86] Monk Nicholas A. M. Maroto, Miguel. *Cellular Oscillatory Mechanisms*, volume 641. Landes Bioscience 1 Springer Science+Business Media, LLC dual imprint1Springer series:Advances in Experimental Medicine and Biology, 2008.

[87] N. J. Martinez, M. C. Ow, M. I. Barrasa, M. Hammell, R. Sequerra, L. Doucette-Stamm, F. P. Roth, V. R. Ambros, and A. J. Walhout. A c. elegans genome-scale microRNA network contains composite feedback motifs with high flux capacity. *Genes and Development*, 22(18):2535–2549, September 2008.

[88] Yoshito Masamizu, Toshiyuki Ohtsuka, Yoshiki Takashima, Hiroki Nagahara, Yoshiko Takenaka, Kenichi Yoshikawa, Hitoshi Okamura, and Ryoichiro Kageyama. Real-time imaging of the somite segmentation clock: Revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1313–1318, 2006.

[89] Mitsuhiro Matsuda, Makito Koga, Eisuke Nishida, and Miki Ebisuya. Synthetic signal propagation through direct cell-cell interaction. *Sci. Signal.*, 5(220):ra31, 2012.

[90] Pankaj Mehta, Ranjan Mukhopadhyay, and Ned S. Wingreen. Exponential sensitivity of noise-driven switching in genetic networks. *Phys. Biol.*, 5:026005+, 2008.

[91] Petros Mina, Mario di Bernardo, Nigel J. Savery, and Krasimira Tsaneva-Atanasova. Modelling emergence of oscillations in communicating bacteria: a structured approach from one to many cells. *Journal of The Royal Society Interface*, (78).

[92] Eric A. Miska, Ezequiel Alvarez-Saavedra, Allison L. Abbott, Nelson C. Lau, Andrew B. Hellman, Shannon M. McGonagle, David P. Bartel, Victor R. Ambros, and H. Robert Horvitz. Most Caenorhabditis elegans microRNAs Are Individually Not Essential for Development or Viability. *PLoS Genet*, 3:e215+, 2007.

[93] Melanie Mitchell. *An Introduction to Genetic Algorithms (Complex Adaptive Systems)*. The MIT Press, February 1998.

[94] Hiroshi Momiji and Nicholas A.M. Monk. Dissecting the dynamics of the hes1 genetic oscillator. *Journal of Theoretical Biology*, 254(4):784 – 798, 2008.

[95] N. A. Monk. Oscillatory expression of hes1, p53, and NF-kappaB driven by transcriptional time delays. *Current Biology*, 13(16):1409–1413, August 2003.

[96] Luis G. Morelli, Sal Ares, Leah Herrgen, Christian Schrter, Frank Jlicher, and Andrew C. Oates. Delayed coupling theory of vertebrate segmentation. *HFSP Journal*, 3(1):55–66, 2009.

[97] Luis G. Morelli, Koichiro Uriu, Saúl Ares, and Andrew C. Oates. Computational approaches to developmental patterning. *Science*, 336(6078):187–191, April 2012.

[98] Shankar Mukherji, Margaret S Ebert, Grace X Y Zheng, John S Tsang, Phillip A Sharp, and Alexander Van Oudenaarden. Microrna s can generate thresholds in target gene expression. *Nat Genet*, 43:854–859, 2011.

[99] Bela Novak and John J. Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, October 2008.

[100] A. C. Oates, L. G. Morelli, and S. Ares. Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development*, 139(4):625–639, January 2012.

[101] Erin L. OBrien, Elizabeth Van Itallie, and Matthew R. Bennett. Modeling synthetic gene oscillators. *Mathematical Biosciences*, 236(1):1 – 15, 2012.

[102] Kathryn A. O'Donnell, Erik A. Wentzel, Karen I. Zeller, Chi V. Dang, and Joshua T. Mendell. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, 435:839–843, June 2005.

[103] Yusuke Okubo, Takeshi Sugawara, Natsumi Abe-Koduka, Jun Kanno, Akatsuki Kimura, and Yumiko Saga. Lfng regulates the synchronized oscillation of the mouse segmentation clock via trans-repression of notch signalling. *Nat Commun. Nature Publishing Group, a division of Macmillan Publishers Limited*, 3(1141), october 2012.

[104] Matteo Osella, Carla Bosia, Davide Cora', and Michele Caselle. The role of incoherent microRNA-mediated feedforward loops in noise buffering. *PLoS Comput Biol*, 7(3):e1001101, 2011.

[105] H.G. Othmer and L.E. Scriven. Instability and dynamic pattern in cellular networks. *Journal of Theoretical Biology*, 32(3):507 – 537, 1971.

[106] Ertugrul M Ozbudak and Julian Lewis. Notch signalling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries. *PLoS Genet*, 4:e15, 2008.

[107] Ertugrul M. Ozbudak, Mukund Thattai, Iren Kurtser, Alan D. Grossman, and Alexander van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature genetics*, 31(1):69–73, 2002.

[108] Louis M. Pecora and Thomas L. Carroll. Master stability functions for synchronized coupled systems. *Physical Review Letters*, 80(10):2109–2112, March 1998.

[109] Michael Petersen and Jesper Wengel. Lna: a versatile tool for therapeutics and genomics. *Trends in Biotechnology*, 21(2):74–81, 2003.

[110] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: A Universal Concept in Nonlinear Science*. Cambridge University Press, January 2002.

[111] Erik Plahte. Pattern formation in discrete cell lattices. *Journal of Mathematical Biology*, 43(5):411–445, 2001.

[112] Athanasios Polynikis, Giulia Cuccato, Stefania Criscuolo, Stephen J. Hogan, Mario di Bernardo, and Diego di Bernardo. Design and con-

struction of a versatile synthetic network for bistable gene expression in mammalian systems. *Journal of Computational Biology*, pages 195–203, 2011.

[113] Olivier Pourquié. Vertebrate segmentation: from cyclic gene networks to scoliosis. *Cell*, 145(5):650–663, May 2011.

[114] Oliver Purcell, Nigel J. Savery, Claire S. Grierson, and Mario di Bernardo. A comparative analysis of synthetic genetic oscillators. *J R Soc Interface*, 7(52):1503 –1524, November 2010.

[115] Arjun Raj, Scott A. Rifkin, Erik Andersen, and Alexander van Oudenaarden. Variability in gene expression underlies incomplete penetrance. *Nature*, 463(7283):913–918, 2010.

[116] Ingmar H. Riedel-Kruse, Claudia Mller, and Andrew C. Oates. Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. *Science*, 317(5846):1911–1915, 2007.

[117] Nitzan Rosenfeld, Michael B Elowitz, and Uri Alon. Negative autoregulation speeds the response times of transcription networks. *J Mol Biol*, 323(5):785–793, November 2002.

[118] Nitzan Rosenfeld, Jonathan W. Young, Uri Alon, Peter S. Swain, and Michael B. Elowitz. Accurate prediction of gene feedback circuit be-

havior from component properties. *Molecular Systems Biology*, 3:143, 2007.

[119] Hiromi Shimojo Itaru Imayoshi Ryoichiro Kageyama, Toshiyuki Ohtsuka. Dynamic notch signaling in neural progenitor cells and a revised view of lateral inhibition. *Nature Neuroscience*, (11):12471251, 2008.

[120] M.H. Jensen S. Pigolotti, S. Krishna. Symbolic dynamics of biological feedback networks. *Phys. Rev. Lett*, 102(118101), 2009.

[121] Michael A. Savageau. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, 252(5484):546–549, December 1974.

[122] O. Shaya and D. Sprinzak. From notch signaling to fine-grained patterning: Modeling meets experiments. *Current Opinion in Genetics & Development*, August 2011.

[123] Hiromi Shimojo, Toshiyuki Ohtsuka, and Ryoichiro Kageyama. Dynamic expression of notch signaling genes in neural stem/progenitor cells. *Frontiers in neuroscience*, 5, 2011.

[124] Velia Siciliano. *A microRNA based genetic clock*. PhD thesis, Open University, 2011.

[125] Velia Siciliano, Immacolata Garzilli, Chiara Fracassi, Stefania Criscuolo, Simona Ventre, and Diego di Bernardo. mirnas confer phe-

notypic robustness to gene networks by suppressing biological noise. *Nature Publishing Group, a division of Macmillan Publishers Limited - Nat Commun*, 4, september 2013.

[126] Velia Siciliano, Filippo Menolascina, Lucia Marucci, Chiara Fracassi, Immacolata Garzilli, Maria Nicoletta Moretti, and Diego di Bernardo. Construction and modelling of an inducible positive feedback loop stably integrated in a mammalian cell-line. *PLoS Comput Biol*, 7:e1002074, 2011.

[127] Sisler HD Siegel MR. Inhibition of protein synthesis in vitro by cycloheximide. *Nature*, 200:675–676, 1963.

[128] A. Singh. Negative feedback through mrna provides the best control of gene-expression noise. *NanoBioscience, IEEE Transactions on*, 10:194 –200, 2011.

[129] Roy S. Smith and John C. Doyle. Model Validation: A Connection Between Robust Control and Identification. *IEEE Transactions on Automatic Control*, 37:942–52, 1992.

[130] David Sprinzak, Amit Lakhanpal, Lauren LeBon, Jordi Garcia-Ojalvo, and Michael B. Elowitz. Mutual inactivation of notch receptors and ligands facilitates developmental patterning. *PLoS Comput Biol*, 7(6):e1002069+, June 2011.

[131] David Sprinzak, Amit Lakhanpal, Lauren LeBon, Leah A. Santat, Michelle E. Fontes, Graham A. Anderson, Jordi Garcia-Ojalvo, and Michael B. Elowitz. Cis-interactions between notch and delta generate mutually exclusive signalling states. *Nature*, 465(7294):86–90, May 2010.

[132] Pasquale Stano. Synthetic biology of minimal living cells: primitive cell models and semi-synthetic cells. *Systems and synthetic biology*, 4(3):149–156, September 2010.

[133] Jesse Stricker, Scott Cookson, Matthew R R. Bennett, William H H. Mather, Lev S S. Tsimring, and Jeff Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221):516–519, October 2008.

[134] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001.

[135] Yannick Sylvestre, Vincent De Guire, Emmanuelle Querido, Utpal K. Mukhopadhyay, Vronique Bourdeau, Franois Major, Gerardo Ferbeyre, and Pascal Chartrand. An e2f/mir-20a autoregulatory feedback loop. *The Journal of Biological Chemistry*, 282(4):2135–2143, 2007.

[136] Zoltan Szallasi, Jörg Stelling, and Vipul Periwal. *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*. The MIT Press, 2006.

[137] Mukund Thattai and Alexander van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.

[138] Mathew J. Thayer, Stephen J. Tapscott, Robert L. Davis, Woodring E. Wright, Andrew B. Lassar, and Harold Weintraub. Positive autoregulation of the myogenic determination gene myod1. *Cell*, 58(2):241–248, July 1989.

[139] Marcel Tigges, Tatiana T. Marquez-Lago, Jorg Stelling, and Martin Fussenegger. A tunable synthetic mammalian oscillator. *Nature*, 457(7227):309–312, 2009.

[140] Marcel Tigges, Tatiana T. Marquez-Lago, Jörg Stelling, and Martin Fussenegger. A tunable synthetic mammalian oscillator. *Nature*, 457(7227):309–312, 2009.

[141] John Tsang, Jun Zhu, and Alexander Van Oudenaarden. Microrna-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell*, 26:753–67, 2007.

[142] Maki Ukai-Tadenuma, Takeya Kasukawa, and Hiroki R. Ueda. Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. *Nature Cell Biology*, 10(10):1154–1163, September 2008.

[143] V.A. Living colors destabilized ecfp and eyfp vectors. *Clontechniques*, XIV(3):14–15, 1999.

[144] Aleksandra M. Walczak, Jos N. Onuchic, and Peter G. Wolynes. Absolute rate theories of epigenetic stability. *Proceedings of the National Academy of Sciences of the United States of America*, 102:18926–18931, 2005.

[145] W Weber, J Stelling, M Rimann, B Keller, M Daoud-El Baba, CC Weber, D Aubel, and Fussenegger M. A synthetic time-delay circuit in mammalian cells and mice. *Proc Natl Acad Sci U S A.*, 104(8):2643–2648, 2007.

[146] Nir Yosef and Aviv Regev. Impulse control: temporal dynamics in gene transcription. *Cell*, 144(6):886–896, March 2011. PMID: 21414481.

[147] Nir Yosef and Aviv Regev. Impulse control: Temporal dynamics in gene transcription. *Cell*, 144:886–896, 2011.

[148] Stefan Zeiser, Johannes Mller, and Volkmar Liebscher. Modeling the hes1 oscillator. *Journal of Computational Biology*, 14(7):984–1000, 2007.